

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 07:55:06 ; Search time 5782 Seconds
(without alignments)
3542.971 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITRYAQTFRGLGCIIT.....PAIPDREVTYRPEDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODX=frame+ p2n model -DEV=xlh
-Q/cg2n2_1/USPTO_spool/US09930591/runat_13092004_164952_700/app_query.fasta.1.839
-DB=EST -QFMT=fastap -SUFIX=fst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09930591 @CGN 1.1 2810 @runat_13092004_164952_700 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEBUFRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_ges_hum:*
18:	em_ges_pln:*
19:	em_ges_vrt:*
20:	em_ges_fun:*
21:	em_ges_mam:*
22:	em_ges_mus:*
23:	em_ges_pro:*
24:	em_ges_pig:*
25:	em_ges_vrl:*
26:	em_ges_vrl:*
27:	em_ges_vrl:*
28:	gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	163	4.5	3291	11	BC046781 Mus muscu
2	150.5	4.2	3229	11	AK004665 Mus muscu
3	141	3.9	3359	29	AY411078 Homo sapi
4	140.5	3.9	822	13	BU054791 UI-M-PD0
5	136.5	3.8	3320	11	BC059369 Homo sapi
6	133.5	3.7	889	14	CD359697 AGENCOURT
7	132.5	3.7	818	13	BU187274 AGENCOURT
8	131	3.6	1201	9	AL560974
9	129	3.6	2972	11	AK031534 Mus muscu
10	129	3.6	2984	11	AK031679 Mus muscu
11	129	3.6	2886	11	AK028274 Mus muscu
12	129	3.6	3956	11	BC043699 Mus muscu
13	128.5	3.6	1697	29	AY404177 Mus muscu
14	127.5	3.5	691	13	BU054966 UI-M-PD0
15	127.5	3.5	738	14	CD240900 AGENCOURT
16	127.5	3.5	2976	29	AY400284 Homo sapi
17	126	3.5	1283	13	BC0709745 AGENCOURT
18	124.5	3.4	623	14	CF131593 UI-M-PD0
19	124.5	3.4	2388	29	AY418898 Mus muscu
20	124.5	3.4	3633	11	AY383690 Rattus no
21	124	3.4	2388	29	AY418896 Homo sapi
22	123.5	3.4	1788	11	AY105041 Zea mays
23	123	3.4	790	14	CK017540 AGENCOURT
24	123	3.4	3461	29	AY398774 Mus muscu
25	123	3.4	3802	11	AK004733 Mus muscu
26	123	3.4	4640	11	BC058331 Mus muscu
27	123	3.4	4640	11	BC062885 Mus muscu
28	122	3.4	3879	11	BC028405 Homo sapi
29	121.5	3.4	3211	11	BC041392 Homo sapi
30	121	3.3	2270	29	AY418897 Homo sapi
31	120.5	3.3	478	10	BF725559 Homo sapi
32	120.5	3.3	694	12	BI088407 Homo sapi
33	120.5	3.3	3856	11	AK078552 Mus muscu
34	120	3.3	1143	12	BM926541 AGENCOURT
35	120	3.3	3494	11	BC013208 Homo sapi
36	119	3.3	919	13	BK424517 Homo sapi
37	119	3.3	1339	14	CF753594 Est-Conti
38	119	3.3	3679	11	AK084541 Mus muscu
39	118.5	3.3	769	9	AU170397
40	118.5	3.3	1008	12	BQ050646 AGENCOURT
41	117.5	3.2	2031	11	BC038434 Homo sapi
42	117.5	3.2	3024	11	AK049321 Mus muscu
43	117	3.2	3008	11	AF318356 Homo sapi
44	117	3.2	3461	29	AY398772 Homo sapi
45	117	3.2	3915	11	BC038417 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS BC046781 3291 bp mRNA linear HTC 19-NOV-2003
DEFINITION Mus musculus DEAH (asp-glu-ala-his) box polypeptide 16, mRNA (cDNA clone IMAGE:5716182), containing frame-shift errors.
ACCESSION BC046781
VERSION BC046781.1 GI:28386173
KEYWORDS
SOURCE HTC.
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (baaes 1 to 3291)

AUTHORS

AUTHORS
Staubsberg, R.L., Feingold, E.A., Grouse, L.H., Berge, J.J.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenn, C.M., Schuler, G.D.,
Altschul, S.P., Zeeberg, B., Buerckow, K.H., Scheffer, C.F., Bat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Donald, M.F., Toshiyuki, S.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrams, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallotton, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Halton, E., Keltman, M., Madan, A., Young, A.C., Stevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Stevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimmwood, J., Schmitt, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalusz, D.E.,
Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

USA was funded by National Cancer Institute, Bethesda, MD 20892-2550, USA. bhargava@nrc.ca

REMARK
COMMENT

Email: cgabos-rcw@uiowa.edu
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA library Arrayed by: The I.M.A.-G.E. Consortium (ULNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonardo, M.F., Akabogu, I., Bail, T., Bail, J., Crouch, R., Davis, A.,
Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein
 This clone has the following problem: frame shifted.
 Location/Qualifiers

FEATURES

```
source
1..3431
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5716182"
/tissue_type="mouse_brain 12.5 dpc"
/clone_id="NH_BMAP_FCO"
/lab_host="DH10B"
/note="vector: ptx-ASC"
```

ORIGIN

Alignment Scores:	
Pred. No.:	0.00271
Score:	163.00
Percent Similarity:	35.15%
Best Local Similarity:	23.87%
Query Matchn:	4.51%
DB:	11
	Gaps: 27
	Length: 3291
	Matches: 127
	Conservative: 60
	Mismatches: 165
	Indels: 181
	Gaps: 27

US-09-930-591-2 (1-686) x BC046781 (1-3291)

[illegible]

QY 396 euApValSerVal11eProthrSerGlyAspValVal1Val1aThrAspAlaLeuM 416
 |||::|
 1895 TCGAGATCCAGCGGACCCCGCCGAGATATACGTGG-----TTCC 1939
 DB
 QY 416 ecThrGlyPheThrGlyAspPheAspSerVal11eAspCys----- 429
 |||::|
 1940 TCACCGGACCA-GGAGAGATTGAGGCTGCTGAGATGCTCCAGAGACCGCTGCCAG 1998
 DB
 QY 429 ----- 429
 DB 1999 CTGGGCTCCAGATCCGGAGACTCTGTGCTGCCATTATGCAACTGCCCTCAGAC 2058
 QY 430 -----AenThrCysValThr- 434
 DB 2059 ATGCAGGCTCGCATCTTCAGGCCACACCCCGGGGGCCGGAAGAACTTCCTCAGCA 2118
 QY 435 --GlnThrValAspPheSerLeuAspProThrPheThr----- 446
 |||::|
 2119 TCGAAGGCATCATCTGTGTGCTGACCCAGGGTTCTCAAGCAGAGAGACTCAACCTC 2178
 QY 447 -----lleglurThrleThrleu---ProGlnAspAlaValSerArgThrGlnArg 463
 |||::|
 2179 GTACGGGAATGGAGTCACTCAGCGTACCCCTGCAGCAGAGGCTTACGCCAATCAGCGG 2238
 DB
 QY 463 rGgYArGThrGlyATGgYlyPrGgYlYlEtyArGpHeValAlaPrGgYlGluArGp 483
 |||::|
 2239 CTGGCGCTCAGGTGCA-----GTGCTGCCGG----- 2267
 DB
 QY 483 roSerGlyMetPheAspSerSerValLeuCysGlnCysTyArAspAlaGlyCysAlaTyr- 502
 |||::|
 2268 -----AAGTGTTCGCGCTGTATACGGCTCGG 2295
 DB
 QY 503 --TyrGluLeuThrProAlaGluThrThrVal 512
 |||::|
 2296 CCTATCAGCATGAGTCAAGAGACCAAGT 2327
 DB

RESULT 2
 AK004665
 LOCUS
 DEFINITION
 AK004665 3229 bp mRNA linear hnc 20-SEP-2003
 Mus musculus adult male lung cDNA, RIKEN full-length enriched
 library, clone:1200009D07 product:PROBABLE ATP-DEPENDENT HELICASE
 DDX35 (DEAH-BOX PROTEIN 35) homolog [Homo sapiens], full insert
 sequence.
 ACCESSION
 AK004665
 KEYWORDS
 HNC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 99279253
 10349636

TITLE
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20493174
 11042159

REFERENCE
 AUTHORS
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861

REFERENCE
 AUTHORS
 4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

TITLE
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 MEDLINE
 PUBMED
 1200009D07

REFERENCE
 AUTHORS
 5 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kaasakawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
 Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submision
 Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research Group, RIKEN Genomic Sciences Center
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail: genome-reseq@riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216]
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGAGCGGCGCGCACTGAGTTTCTTTTCTTTTCTT 3'], cDNA was
 prepared by using triazole thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adaptor of sequence [5'
 GAGAGAGAGAGATCCAGATCCAGATCAATTAATTAATTAACCCCCCCCC 3']. cDNA was
 cleaved with XhoI and SacI. Cloning sites, 5' end: SacI; 3' end:
 XhoI. Host: SOLR.

FEATURES
 source
 1. 3229
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:1200009D07"
 /db_xref="MG1:1904330"
 /db_xref="taxon:10090"
 /clone="1200009D07"
 /sex="male"
 /tissue_type="lung"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1. 3229
 /note="PROBABLE ATP-DEPENDENT HELICASE DDX35 (DEAH-BOX
 PROTEIN 35) homolog [Homo sapiens] (SWISSPROT Q9H521,
 evidence: FASTA, 94.2%id, 99.8%length, match=2126)"

misc_feature
 1. 3229

ORIGIN
 Alignment Scores:
 Pted. No.:
 Score: 0.0368 Length: 3229
 150.50 Matches: 121

Percent Similarity:	34.22%	Conservative:	72
Best Local Similarity:	21.45%	Mismatches:	221
Query Match:	4.16%	Indels:	151
DB:	11	Gaps:	24
US-09-930-591-2 (1-686) x AK004665 (1-3229)			
QY	126 SerLeuLeuSerPro--ArgProIleSerTyLeuLysGlySerSerGlyProLeu	144	
DB	6 ACCATGGCTGGCCCTTGGAGCAGTAATTCGGAAACCGGATACAAAGGGCCA---	62	
QY	145 LeuCyProAlaGlyHISAlaValGlyIlePheArgAlaAlaValCysThrArgGlyAla	164	
DB	63 -----GGTCTCAACATTTCTGAAAGAGCAAAAGTCT---ACCGAAGACTCT	107	
QY	165 AlaAlaValAlaAspPheIleProValGlySerLeu-----GluThrThrMetArg	181	
DB	108 GGGACAAACCATATTACAAATCCCTACGCTGCTCTATAGAGCAGCAGCGGCGAAG	167	
QY	182 SerProValPheSerAspSerSerProProAlaValProGlnSerTyGlnValAla	201	
DB	168 CTGCCCGTGTCAAGCTTGAACACATTTTGTACTGTGTAGAAACATATACAGCCGTG	227	
QY	202 HISLeuHISAlaProThrGlySerGlyLysSerThrLysValPro-----	216	
DB	228 GTGATTTGGTGAAGAAACAGATGTGGAAAGAGCACCAGATTCACAGTACCTGGCAGAA	287	
QY	217 AlaAlaTyraAlaIleGlnGlyTyTyLysValLeuValLeuAsnProSerValAlaAlaThr	236	
DB	288 GCTGGCTGACAGCAGGAGAGCAGTGTGCGAGTACCCAGCTCGCAGAGTGGCTGCC	347	
QY	237 MetGlyPheGlyAlaTyMetSerLysValHisGlyIleAspProAsnIleArgThrGly	256	
DB	348 GTGACGCAATGA-TCCTTCTTCCAAAGGTTCACAGCCGAGTACGTATGAAGGGGTGA	406	
QY	257 ValArgThrIleThrThrGlySerProIleThrTy-----Ser	269	
DB	407 GTGCTGGCGCATGAAGTGGCTACTGTATCGCTGCATGACTGCACCGACCATGGCC	466	
QY	270 ThrTyGlyLysPheLeuAlaAspGlyGlyCysSerGly-----	282	
DB	467 ACCAGATCAAGATTCCTTACAGATGGCATGCTGTCAGAGAAATGATGGTCCGCTG	526	
QY	283 ---GlyAlaTyraAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSerIle	301	
DB	527 TTAAACAAATATATGATCATCATGCTGATGAAGCCACAGAGAGACCTTGACAGGAC	586	
QY	302 LeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeu	321	
DB	587 ATGGCCATGTGGCTGTGAAGAGATTCAGAAAAGCGAGGGGATCTTCGCTGATTTGTG	646	
QY	322 AlaThrLalaThrProProGlySerValThrValProHisProAsnIleGluGlu----	339	
DB	647 GCCTCAGCCACTGTGACGAGAGAAATTCGAGATTTCTTTAACCAAGATGAGACCACT	706	
QY	340 -----ValAlaLeuSerThrThrGlyGlu-----	347	
DB	707 GACCCAGCAGAGATACCTCTGTGACCTCAGGTGAGAGAGACGACATTTCCAGTGGAT	766	
QY	348 -----IleProPheTyGlyLysAlaIleProLeuGluAlaIle	360	
DB	767 ATCTTTTACCTACAAAGTCCGTGTCAGATTAATATCAAGACACTGTGACACCGTGTA	826	
QY	361 Lys-----GlyGlyArgHisIlePhePheCysHisSerTyLysLys	374	
DB	827 AAAATTCATCAGACAGAGAGATGAGACATTAAGCTTTTCTTACTGGCCAGAGAA	886	
QY	375 CysAspGluLeuAlaAlaLysLeuVal-----AlaLeuGly-----	386	
DB	887 GTAAGACTGTGTGTGCATGTATGATGAGCAGCCGCGGGCTGGCTCGCACTGGGATG	946	
QY	387 -----ValAlaAlaValAlaLysTyraArgGlyLeuAsp-----	397	

DB	947 AAGAAACACTCCGGGTCTCCCATGATATGACAGACTGCTTTTGGACAGATGAAG	1006	
QY	398 ---ValSerValIleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMet	416	
DB	1007 GTGTTTGAAGAGGTGCACACAGGCTCAGAGAGTATCGTGGCCACCAATGTGGCAGAG	1066	
QY	417 ThrGlyPheThr---GlyAspPheAspSerValIleAspCysAsnThrCysValThrGln	435	
DB	1067 ACTTCATACACAAATCAGAGGATGTGTATGTGATTAAGTCTGTGGCTTTATAGACTGCA	1126	
QY	436 ThrValAspPheSerLeuAspProThrPheThrIleGlu---ThrIleThrLeuProGln	454	
DB	1127 -----GCTCAACACCCAGACACACTATGTATGACTGTGGTGGTGTACCAATG	1174	
QY	455 AspAlaValSerArgThrGlnArgArgGlyTyraThrGlyArgGlyLysProGly---Ile	473	
DB	1175 TCTCAGGGCTAGCCAAATCAGCGGCGAGAGTGGTGGCCCAACCGCTCGGAAAGTGT	1234	
QY	474 TyraArgPheValAlaProGlyGlyLysProSerGlyMetPheAspSerSerValLeuGly	493	
DB	1235 TATGCTCTTACACA-----GAGGAGCCTTTTGACACGACTACCTCACTCCAGCTCCT	1288	
QY	494 GlyCysTyraAspAlaGlyCysAlaIleTyTyGlyLeuThrProAlaGluThrThrValArg	513	
DB	1289 GAGATGCAAGCGCAGCAATTTG-----GCCCGCTATCTCGCAG	1327	
QY	514 LeuArgAlaTyMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPheThr	533	
DB	1328 CTAAAGCC-----	1336	
QY	534 GluGlyValPheThrGlyLeuThrHisIle---AspAlaHisPheLeuSerGlnThrLys	552	
DB	1337 -----CTAGGGATACAAATGATGCTCAGAGTTCACCTTCAATGATGCT-----	1375	
QY	553 GlnSerGlyGluAsnLeuProTyLeuValAlaTyGlnAlaThrValCysAlaArgAla	572	
DB	1375 -----	1375	
QY	573 GlnAlaProProSerThrAspGlnMetTyLysCysLeuIleArgLeuLysProThr	592	
DB	1376 -----CCCCGCCAGCA---CAGTGCATGGTTCAGGCTTTG-----	1408	
QY	593 LeuHisGlyProThrProLeuLeuTyraArgLeuGlyAlaValGlnAsnGluValThrLeu	612	
DB	1409 -----GAGCTGCTTAATGCTCTCGAGAGCTTCGACAAAGACTGTGGCCTTA	1453	
QY	613 ThrHisProVal	616	
DB	1454 ACTGAGCTCTT	1465	
RESULT 3			
AY411078 3359 bp DNA linear GSS 16-DEC-2003			
LOCUS Homo sapiens TIE gene, VIRUAL TRANSCRIPT, partial sequence,			
DEFINITION genomic survey sequence.			
ACCESSION AY411078			
VERSION AY411078.1 GI:39767046			
KEYWORDS GSS.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
REFERENCE 1 (bases 1 to 3359)			
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,			
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,			
Adams,M.D. and Carcill,M.			
Inferring nonneutral evolution from human-chimp-mouse orthologous			
gene trios			
JOURNAL Science 302 (5652), 1960-1963 (2003)			
PUBMED 14671302			
REFERENCE 2 (bases 1 to 3359)			
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,			
AUTHORS			

Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
Location/Qualifiers
1..3359

source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3359
/gene="TIE"
/locus_tag="HCM4104"

ORIGIN

Alignment Scores:

Pred. No.:	0.323	Length:	3359
Score:	141.00	Matches:	168
Percent Similarity:	31.38%	Conservative:	84
Best Local Similarity:	20.92%	Mismatches:	233
Query Match:	3.90%	Indels:	318
DB:	29	Gaps:	41

US-09-930-591-2 (1-686) x AY411078 (1-3359)

```

QY 39 ThrAlAlGlnThrPheLeuAlaThrCysGleAsnGlyValCysTrpThrValTyRHis 58
DB 42 ACGAGACCCCGAGCGCTTCTCCAGACTGCGTGTGGAGAGGC----- 86
QY 59 GYAlAGly----- 63
DB 87 GGGGGGGAGGGGGCTCGAGCGCTGGGGCCCGCTGCTGTGAGAAAGACGACGCT 146
QY 64 ThrIleAlaSerProlysglyProvalIleGlnMetTyThrAsnValAspGlnAsp--- 82
DB 147 ATGTGCGGACCCCGCGCGGCCACCCCTGCGCTGGCGGCAACGGTTGCGACAGGTC 206
QY 83 --LeuValGlyTrp-ProAlaPro----- 89
DB 207 ACGCTTCGGCGGCTTCTCCAGACCCCTCGACCTGCGGCGCTTCTCTGCGGCGGT 266
QY 90 -----GlnGlyAlaArgSerLeuThrProCysThrCysGlySerSerAspLeuTyR 107
DB 267 GCTGGGGCGGGCGGCGCGGCTGATCTTACGTGACAAACGCCCTGAGCCACCTGCTT 326
QY 107 uValThrArgHis-----AlaAspValIleProVal----- 117
DB 327 CCAAGACAGGTGACACACACTGTGAACAAAGGTGACACCGCTGACTTCTGACCGTGG 386
QY 118 -----ArgArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIle 133
DB 387 CACAAAGAGAAAGACAGACAGCTGATCTGGAAAGAGACAGATCTCACTTACACCTG 446
QY 133 eSerTyRLeuTyGlySerSerGlyTyRProLeuLeuCysProAlaGlyHisIleAlaGle 153
DB 447 GACTGCAATGAACCCACGATGGGCGGTTCC-----TGCTGCAAGTCCCAAAATGGCGAG 500
QY 153 yIlePheArgAlaAlaValCysThrArgGlyValAlaIleValAlaAspPheIleProVa 173
DB 501 CCAACATCGAGCGGCA-----TCATCACTGCACT 530
QY 173 lGlnSerLeuGlnThrThrMetArgSerProValPheSerAspAsn----- 188
DB 531 TACCTGAAAGCCAGCCCTGCGGACAGCGCTTCTTTCGCTCATCTGCGGCGGTTGGG 590
QY 189 -----SerSerProAlaValProGlnSerTyRValAlaHisIleu----- 203
DB 591 GCTGGCGCTGGGGCGGCGGCTGTATCCAAAGAGTCCCAAGTTGACATGGAAGGTTC 650
QY 204 -----HisAlaProThrGlySer---GlyLysSerThrTyR 214

```

```

DB 651 TCCCAAGACCATGACGCGGATGTATGTATGCCCCCTGGCTTCACTGGACACCGCGTGGAA 710
QY 214 eValProAlaAlaTyRAla----- 225
DB 711 CAGGCTTGAGAGAGAGGCGCTTTGGGACAGACTGCCAGAGACAGTCCAGCATATCA 770
QY 225 eValLeuValLeuAsnProSer-----ValAlaAlaThrMetGlyPheGle 240
DB 771 GGCTGCGGGGCGCTCATCTTGTGCTCCAGACCCCTATGGCTGTGCTGTGGATGTGCG 830
QY 240 yAlaTyRMetSerTyRAlaHisGlyTyRAspProAsnIleArgThrValAlaArgThrI 260
DB 831 TGGAGAGAGAGCGAGTCCCAAGAGCTTGTGCTCCCTGTCATTTGGGG----- 879
QY 260 eThrThrGlySerProIleThrTyRSerThrTyRgLyPheLeuAlaAspGlyGlyCys 280
DB 879 ----- 879
QY 280 sSerGlyGlyAlaTyRAspIleIleCysAspGlyCysHisSerThrAspAlaThrSe 300
DB 880 -----CTGATTCGCGACTTCAGTGCAGTGCAGATGTGCAGATGTGCG 917
QY 300 rIleLeuGlyIleGlyThrValLeuAspGlnIleGlnThrAlaGlyAlaArgLeuTyR 320
DB 918 ACTTGTG--ACCGGTTCAAGTGTGTGTGTGCTCCCTCTGGGTGCATGAGTGCACGT 974
QY 320 lIleuAla---ThrAlaThrPro----- 326
DB 975 GAGAGTCAAGACCGGATCCCAAGATCTCAACATGCGCTCAAGATCGATCAACTTA 1034
QY 327 -----ProGlySerValThr-----ValProHisPro----- 335
DB 1035 GAGAGATGCCCGGATCAACGTGACGTGACGAGGAACCCCTTCCCGTGGGGGCGAGC 1094
QY 336 -----AsnIleGlnGlnValAlaLeuSerThrThrGlyGlnIleProPheTyRgLy 353
DB 1095 ATGAGCTACGAGACAGACAGCGCACTGTGCTGTCCACCAAGGCACTTGTGGAGCA 1154
QY 353 sAlaIlePro-----LeuGlnAlaIleLeuGly 363
DB 1155 GAGAAACCAACACTGATTCAGAGTCCCGCTGTGTTCTTGGGACAGTGGCTTCTGG 1214
QY 363 yArgHisLeuIlePheCysHisSerTyRLeuTyRAspGlnLeuAlaIleLeuVal 383
DB 1215 GAGTGGCGG-----TGTCCACATCTGGCGGCAAGCACG 1250
QY 383 l-AlaLeu-GlyValAsnAlaValAlaTyTyRArgGlyLeuAspValSerValIlePro 402
DB 1251 CGGCGCTCAAGTCAAT-----GTGAAGTGGCCCCC 1283
QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThr----- 417
DB 1284 -----GTGCCCTGTGTCACCTCGGCTTCTTCAACAGAGAGCGGCCAG 1328
QY 418 -----GlyPheThrGlyAsp----- 422
DB 1329 CTGTGTGTCTCCCGCTGTGTCTGTCTGTGGAGTGAACCATCTCCAGTCCGCGTGG 1388
QY 423 -----PheAspSerValIleAspCysAsnThrCysValThrGlnThrValAsp 438
DB 1389 CACTACCGGCGCCAGAGACATGACATGACTGTGACATCATTTGTG----- 1433
QY 439 PheSerLeuAspProThrThrThrIleGlnThrIleLeu-----ProGln 454
DB 1434 -----GTGGACCCAGT-----GAGAACGTGACCTTAATGAACCTGAGGCCAAAG 1478
QY 455 AspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyR 474
DB 1479 ACAAGATACAGTGTCTGTGTGACGCTGAGCCGCGGAGGAGAGCA----- 1523
QY 475 ArgPheValAlaProGlyGln-ArgProSerGlyMetPheAspSer-----SerVa 491

```


Qy	218	AlaTyrAlaIaaInglIyTyrLysValLeuValLeuAenProSerValAlaIaIaThreMet	237
Db	164	GGTATATACAAACAAGGGATATGAAGATTCCCTGCACCCCAACCCCGAGAGTGGCTGCATG	223
Qy	238	GIyPheGIaIaTyrMetSerLysAlaHisGIy-----IleAspPro	251
Db	224	AGTGTGGCCGCCGAGTGGCCCGGAGATGGTGTGAAGCTTTGGGAATAGAGTTGGCTAC	283
Qy	252	AsnIleArg-----ThrGIyValArgThrIleThrIleThrGIySerProIleThr	267
Db	284	AGCATCCGCTTTGAGAGACTGCACATCAAGAGGAATCTGC-----CTCCGC	328
Qy	268	TyrSerThrTyrGIy-----LysPheLeuAlaAspGIyGIyCysSerGIyGIy	283
Db	329	TACATGACAGATGGAGATCTCTCCGGAGTTCCTCTTGAGCCTGACCTGAGC-----	382
Qy	284	AlaTyrAspIleIleIleCysAspGIuCysHisSerThrAspAlaThrSerIleLeuGIy	303
Db	383	AGTTCACGCGGTGATGTGTGATGAGCACAACGAAGACCTTACACAGACATCTCTC	442
Qy	304	IIeGIyThrValLeuAspGIaIaGIuThrAlaGIyAlaArgLeuThrValLeuAlaThr	323
Db	443	TTTGGATTGATCAAGAGATGTGTGCTCGCTCCGACCTGACCTCAAGAGTCTGTGGCTTCA	502
Qy	324	AlaIaThrProProGIySerValThrValProHisProAsnIIeGIuGIuValAlaLeuSer	343
Db	503	GCCACA-----ATGACACATGCCCGCTTTTTC	529
Qy	344	ThrThrGIyGIuIleProPheTyrGIyLysAlaIleProLeuGIuAlaIleLysGIyGIy	363
Db	530	ACC-----TTCTTTGATGACGCC--CTGTGTTTGGAAATCCCGGACGC	571
Qy	364	ArgHisLeu-----IlePheCysHisSerLysLysCysAspGIuLeuAlaLys	381
Db	572	AGGTTTCTCTGGACATCTTCTACACCAAGGCTCAGAGGCTGACTACTTGAAGCTTGT	631
Qy	382	LeuValAlaLeuGIyValAsnAlaValAlaIaTyrTyrArgGIyLeuAspValSerValIle	401
Db	632	GTAGATCTCTG-----TTGCAGATCCATGTGACC	661
Qy	402	ProThrSerGIyAspValValValAlaIaThr	412
Db	662	CAGCCCCCTGGGATATCTGTGTCTGCTGACA	694
RESULT 7			
LOCUS	BUI187274	818 bp	mRNA linear EST 04-SEP-2002
DEFINITION	AGENCOURT_7825134 NIH_MGC_67 Homo sapiens CDNA clone IMAGE:6142792		
ACCESSION	BUI187274		
VERSION	BUI187274.1	GI:22701258	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 818)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@db-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LLM13465 row: c column: 17 High quality sequence atop: 546. Location/Qualifiers 1..818		
FEATURES			
Source			

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6142792"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH MG_C 67"
/note="Organ: eye; Vector: pCMV-SPOB6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

```

Alignment Scores:		Pred. No.:	0.169	Length:	818
Score:	132.50	Matches:	63		
Percent Similarity:	41.67%	Conservative:	37		
Best Local Similarity:	26.25%	Mismatches:	94		
Query Match:	3.66%	Indels:	46		
DB:	13	Gaps:	13		
US-09-930-591-2 (1-686) x BU187274 (1-818)					
QY	128	LeuSerProArgProIleSerTyrLeuIleGlySerSerGlyProLeuLeuCyPro	147		
DB	50	STTCGAGAGCCCAAGATCACTGCTGCTGAGAGAGAGACCAATGATGATTCGG	109		
QY	148	AlaGlyHisAlaVal-----GlyIlePheArgAlaIa-ValCysThrArgI	163		
DB	110	---GGCACTCAGCTCCAGGGTATGAGAGCCCTCAGCCCACTTCACTCAGGC	166		
QY	163	ValAlaIaLeuAlaValAspPheIleProValIleGlySerLeuGlyThrThrmcArgSer-	182		
DB	167	CCACCAAAA-----GAGTCATCAAGCCGTCGCGCAGACCT	205		
QY	183	-ProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnValAlaI	202		
DB	206	CCCGGTGCTCCATTTGAGAGAGAGCTCTGGCTGCTATTGCAAAATCAGCAAGCTCTCAT	265		
QY	202	AlaIaIaLeuProThrIleSerGlyIleSerThrIleValPro-----Al	217		
DB	266	CATTGAAGCGGACAGAGCTCAGGAAAGACCAACCAATGCCGAGATCTCTTGAGGA	325		
QY	217	AlaIaTyrAlaAlaGlnIleTyrIleValIleValIleAsnProSerValAlaIaIaThrMe	237		
DB	326	GGGTATATACAAACAAGGATGAAGATGGCTGACCCAAACCCGAGAGTGGCTGCAT	385		
QY	237	IleGlyPheGlyAlaIaTyrMetSerIleValAlaIleGly-----IleAspPr	251		
DB	386	GAGGTGGCCGCCCAAGTGGCCGCGGAGATGGGTGTGAAACCTTGGAATGATGGTGGCTA	445		
QY	251	AsnIleIaArg-----ThrGlyValArgThrIleIleThrThrGlySerProIleTh	267		
DB	446	CAGATCCGCTTGAGACATGCACATCAGACGCAACTGTC-----CTCCG	490		
QY	267	TyrSerThrTyrIleGly-----LysPheLeuAlaAspGlyGlyCysSerGlyI	283		
DB	491	CTACATAGACAGATGGAGATGCTTCCGGAGTTCCTCTGAGAGCTGACCTGGCG----	545		
QY	283	ValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaIleSerIleLeuG	303		
DB	546	-AGTTACAGCGGTGGTGTGATGGATGGAGCAACGAAGAACCCTACACACAGACATTC	604		
QY	303	YleIleGlyThrValLeuAspGlnAlaGlyIleThrAlaGlyAlaArgPheThrValIleuAlaTh	323		
DB	605	CTTTGAGATGATCAAGATGTGCTGCTGCTCCGA--CTGAGCTCAAGGCTCTGGATGG	661		
QY	323	IaIaIaThrProProGlySerValThrValPro-----HisProAsnIleGlnIu	339		
DB	662	CTTCAGCCACAAATGG-----ACACTGCCCCGTTTTCACACCTCTTTGATGA	713		

LOCUS AL560974 1201 bp mRNA linear EST 31-MAY-2003
 DEFINITION AL560974 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CS0DL006YA04 5-PRIME, mRNA sequence.
 ACCESSION AL560974
 VERSION AL560974.2 GI:31285103
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1201)
 REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12907950.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3796.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DL006BA02QPI&cluster=3796.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DL006BA02QPI.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL006YA04"
 /cell_line="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (4T) primer. Five prime and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.:	0.458	Length:	1201
Score:	131.00	Matches:	72
Percent Similarity:	37.11%	Conservative:	46
Best local Similarity:	22.64%	Mismatches:	109
Query Match:	3.62%	Indels:	92
DB:	9	Gaps:	14

US-09-930-591-2 (1-686) x AL560974 (1-1201)

QY 128 leuSerProArgProIleSerGlyLeuGlySerGlyGlyProLeuLeuCyPro 147
 |||||
 123 CTCTCAGAGGCCCACTGATCACTGCTGAGAGGAGGAGACCATGAGTTCG 182
 |||||
 QY 148 AlaGlyHisAlaVal-----GlyIlePheArgAlaAla-ValCysThrArgG1 163
 |||||
 183 ---GGCCACTCAGCTCCAGGGGATGAGAGAGCCGCTCAGCTCACCACCTTCACTCAGGC 239
 |||||
 QY 163 yValAlaIysAlaValAspPheIleProValGluSerLeuGluThrThrMetArgSer-- 182
 |||||
 240 CCAGCAGAAA-----GAGTCATCCAGCGCGCTCCGCCGACGCT 278
 |||||
 QY 183 -ProValPheSerAspAsnSerSerProProAlaValProGlnSerGlyGlnValAlaHi 202
 |||||
 Db 279 CCGCGGTTCCATTTCAGAGAGAGCTCCTGCTCTATTGCAATTCACCAAGTCTCAT 338
 |||||
 QY 202 sleuHisAlaProThnGlySerGlyLysSerThrLysValPro-----Al 217
 |||||
 Db 339 CATTGAAGCGCAGACAGGCTCAGGAAAGACACCCAGATCCCGCAGTATCTCTTGAGGA 398

QY 217 aAlaTyraAlaIaGlnGlyTyTyLysValLeuValLeuAsnProSerValAlaAlaThrMe 237
 |||||
 Db 399 GGGTTATACAAACAAGGTTATGAGATGCTCCGACCCCAACCCGAGAGTGGCTGCAT 458
 |||||
 QY 227 tGlyPheGlyAlaTyraMetSerLysAlaHiGly-----IleArgPr 251
 |||||
 Db 459 GAGTGTGCGCCGCCGAGTGGCCGGAGATGGTGTGAAGCTTGGGAATGAGTGGCTGA 518
 |||||
 QY 251 oAnIleArg-----ThrGlyValArgThrIleThrThrGlySerProIleTh 267
 |||||
 Db 519 CAGCATCCGCTTTAGAGACTGCACATCAGAGAACTCTC-----CTCCG 563
 |||||
 QY 267 rTySerThrTyrgly-----LysPheLeuAlaSerGlyGlyCysSerGlyG1 283
 |||||
 Db 564 CTACATGACAGATGGATGCTCTCCGGAGATCTCTCTGCTGCTGACCTGAC 618
 |||||
 QY 283 yAlaTyraPheIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuG1 303
 |||||
 Db 619 -AGTTACAGCGTGTGATGTGATGAGACACACAAAGACCTTACACACAGATTC 677
 |||||
 QY 303 yIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrValLeuAlaTh 323
 |||||
 Db 678 CTTCGATTGATCAGAGATGTTGCTCGCTTC----- 708
 |||||
 QY 323 rAlaThrProProGlySerValThrValProHisProAnIleGluGluValAlaLeuSe 343
 |||||
 Db 709 -----CGACCTGATCTCAAGTCTGTGGCTTC 737
 |||||
 QY 343 rThr-ThrGlyGluIleProPheTy----- 351
 |||||
 Db 738 AGCAATGACACATGCGCCGCTTTTCCACTTCTTGAAGAGCCCTGTTGGAATCC 797
 |||||
 QY 352 --GlyLysAlaIleProLeuGlnAlaIleLysGlyGlyArgHisIleuIlePheCysHis 371
 |||||
 Db 798 CCGAGACGACGTTCTCTGTGAC-----ACTTCTACACCA 833
 |||||
 QY 371 erLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyValaAsnAlaVala 391
 |||||
 Db 834 AGGCTCCAGAGGCTGACTTGAAGCTTGTGATGATCTGTG----- 877
 |||||
 QY 391 laTyTyraArgGlyLeuAspValSerValIleProThrSerGlyAspVal 407
 |||||
 Db 878 -----TTGCAGATCCATGTGAC-CCACCCCTGGGGATATC 912
 |||||

RESULT 9
 AK031534
 LOCUS AK031534 2972 bp mRNA linear HTC 18-SEP-2003
 DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
 enriched library, clone:603048F20 product:hypothetical P-loop
 containing nucleotide triphosphate hydrolase structure containing
 protein, full insert sequence.
 ACCESSION AK031534
 VERSION AK031534.1 GI:26327388
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2
 REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3

AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komori, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hata, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Ozawa, M., Ohara, E., Watanabe, M., Okazaki, Y., Ishikawa, T., Tanaka, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
20530913
11076861

REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS
6
(bases 1 to 2972)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Furuta, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kaya, H., Kawai, J., Kojima, Y., Komodo, S., Komori, H., Kouda, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 2972
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:6030448F20"
/db_xref="WGI:2394772"
/db_xref="taxon:10090"
/clone="6030448F20"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
17. 2641
/note="unamed protein product; hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC27438.1"
/db_xref="GI:26327389"

FEATURES
source

CDS

/translation="MAAGRVGPGPSPRPPTMAEPWKKRLRRRTGASQGRSDSDGSE
FEIQAEADAAKRGKGRALRPSFTSCVSDVEDEMTVAQMKKKSGSQSMGL
SYPEKGIIMKGVVPTPIORITPLVLDKDVAMRTSGKTACFLPMEFLKAR
SAQCRGALLISPTREILKQETKEKTKGRTGKTLIIIGDGMEDQALHNPD
IATPFGLVHVAENMILKQSEVYVDEADRLTEMPGABELOQITIGLPBGHDTVL
SATLPKLVEPAQGLTEPLIRLDVDSKLEQIKTFFLVREDTKAVLLYQNV
RPOQTVPAATKHAAYLTLELMQGVSCAHYISALDQTKNLKFTNKSTLL
VTDLAAGLPIPLDNIYNTSPAKGFLPHRVRVARAGSGTAVSLVADDEVYLL
DLHLFLRSVTIARCEPSVADAVGDGVGVGVSDVDEDSLOTAMGASLDLG
LHRVANAQOOVVRSPAPSPESIKRAKEILDIAELGHPLESCFEEGELQRLRVD
IKYRTTITTEIYIASSKDSQSMKRAKQDRDAVAFQQRBERQRPDAPQRE
LPQEEEMVEVTEGVTVEVQQRPRPQSGAKRRMETRQDQETVYRPFQD
SERGLSVSGAGAFQOQVAVLDLMDQENMSRQOQLMDKRRKRFVQSQEDK
KKIKTESGRTFSSYKRDLYQMKQKQIDRDNDSRQPSRQRPSPRGKRGSSQ
TSQRPASSVPAGRMRSKLTKEQILKQRPQKQRFQRGGLKQLSANRRRAQELRQ
GARGRGAPSRKGRKRM"

polyA_signal
2954..2959
/note="putative"
2972
/note="putative"

polyA_site
2972
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 3.47 Length: 2972
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK031534 (1-2972)

QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrGlnAlaAspVal 114
Db 90 TGGCGAAACGCGGAACGTGGGCTTCCAGAGCCCGACAGGAGCTCGATAGCGCGGACT 149
115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db 150 TCGAGATCCAGGCGGAGATGACCCCGGAGAGAGAGCTGGCGCTCGAGAGCTTGC 209
QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyProLeuLe 145
Db 210 CCTGATTTCCATCCCTCAGAGCTGATGAGTGCAGCCGACATCGGAGATGTGTGC 269
QY 145 uCyProAlaGlyIleAlaValGlyIlePheArgAlaVal----- 159
Db 270 GAGCCGAGAACAGAAAGAGAGTGTGAGGCTTCCAGTCCATGAGGCTGAGTTACC 329
QY 160 -CysThr-ArgGlyValAlaValAlaValAspPheIleProValGlySerLeuGluThr 179
Db 330 CTGCTTCAAGGAGATGATAAAGGCTTCAAGGTGCCAGGCCATCCAGAGAGAGA 389
QY 179 hrMetArgSerProValPheSerArgPheAsnSerSerProProAlaValProGlnSerTyrG 199
Db 390 CCATC-----CCGATATCTTGATGCGCAAGATGTGTGGCCATG----- 430
QY 199 InvAlaIleIleHisAlaLeuProThrGlySerGlyLysSerThr----- 213
Db 431 -----GCCGAGACAGGCAAGTGGCAAGAGCGGCTGCTCTCTCCCGGA 473
QY 214 -----LysValProAlaIleAlaIleAlaGlnGlyTyrGlyValIleValLeuAsp 231
Db 474 TGTTTGAGCGGCTGAAGGACGAGCTGACAGAGGGGGCTCGAGCCCTCATCTCTCAC 533
QY 231 roSer-----ValAlaIleThrMetGlyPheGlyValAlaTyrMetSerLysAlaHis 248
Db 534 CCACCCGAGAGCTGGCCCTGACAGACATGAAGTTCATTAAGAGCTAGGCAAGTT-CACC 592
QY 248 L-----IleAspProAsnIleArgThrGlyValArgThrIleThrGlyS 264
Db 593 GGCCTCAAGACTGCCTTGATCTCGGTGAGACAAATGAAAGCAGTTTGACGCCCTG 652
QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284

Db	653	CACGAGAACCTGTGACATATATCATTTGGCCACCCTGGGGGTCTGGCATGT-----GGCT	706
Qy	284	LaTyraaPlelleleleCysaaspGlyCysHisSerThrAspAlaThrSerIleuGlyI	304
Db	707	GTGGAGATGAACCTTGAAGCTGCAGAGTGT-----GAGTATGTGTG	748
Qy	304	Le-----	304
Db	749	TTTCATGAGACAGACAGGCTCTTTGAAATGGCTTTTGTAGACAGTACAGAGATCAT	808
Qy	305	-----GlyThrValIleuAsp-----GlnAlaGlyThrIaG	315
Db	809	GGCGGCTTCTGGGGGCCACCAAGAGGTGTCTTTCAGTACATCTGCCCAAGTGTCTG	868
Qy	315	lyAlaArgLeuThr-----ValIleuAlaThrAlaThrProProGlySerValThrV	332
Db	869	GTGGAATTTGGACACAGGACGAGCTCAGAGCCCGTGCATCCCGCTGGACGTAGATCC	928
Qy	332	aIProHisPProAsnIleGluIleuAlaIleuSerThrThiGlnIlePro-----	349
Db	929	AAGTCATAGACAGCTTACAGACTCT-----CTTCTCTTGTG	967
Qy	350	-----PheTyGlyLysAlaIleProleuGlnAlaIleLys-----GlyG	363
Db	968	CCGGAAGACACCAAGGCTGTGTGTCTCTTACCTGTGCAGATGTCTTGGCCGACG	1027
Qy	363	ly-ArgHisleuIlePheCysHisSerIleLysIleCysAspGluIleAlaIleLysLeu	382
Db	1028	GACCCAGACTGTGGTGTGTGTGACCAACAAAGACCATCGGAGTACCTCAGACAGTGTG	1087
Qy	383	ValAlaIleuGlyValAsnAlaValAlaIaTyTrArgGlyLeuAspValSerValIlePro	402
Db	1088	ATGGGCCACAGGCTGTAGTTGGCCCAACATATAGTGTGGACACAGC-----	1138
Qy	403	ThrSerGlyAspValValValAlaThrAspAlaIleuMetThrGlyPheThrGlyAsp	422
Db	1139	-----GCCCGAAGATCAACTGGCCAGCTTACACACAC	1174
Qy	423	PheAspSerValIleAspCysAsnThr-----CysValIleThrGlnIleThrValAspPheSerLeu	441
Db	1175	-----AAATGTTCAACCTTCATCTGATCTGACTGCTGGCCGCCGGGGCTG	1219
Qy	442	AspProThrPheThrIleGluThrIleThrIleuProGlnAspAla-----	456
Db	1220	GACATCCCACTGTGGACAAGCTCACTCACTACACTTCTCCCTGCCAAGGCAAGCTCTTC	1279
Qy	457	-----ValSerArgThrGlnIleArgAsGlyArgThrGlnIleArgGlyLysProGlyIle	473
Db	1280	CTGCACCGAGTGGCGCTGTGGCCGAGCAAGGCCAAAGTGGGACAGC-----	132
Qy	474	TYrArgPheValAlaIleProGlyLysIleArgProSerGlyMetPheAspSerValLeuCy	493
Db	1328	TATCTCTTGTGGGCCCAAGCAGAGTCCCTTACC-----	1361
Qy	493	seGlyCysTyTrAspAlaIleGlyCysAlaIleTrpTyGluIleuThrProAlaGluThrThrValAr	513
Db	1362	-----TCTTGAACCTACACTGTCTTCGGGCGCGCTGTGACCTGGCCGCTTGTGGAG	141
Qy	513	G---LeuAspAlaIleuMetAsnThrProGlyLeuProValCysGlnAspHisIleuGlnPh	532
Db	1418	GAGCTTCACTGAGCAATGCGGTGGCAAGG-----ACGGAATG	145
Qy	532	eTrpGlyLys-----	535
Db	1457	CTGGGTGCGCTGGCCCAAGTGTATGATGATGAGACAGACAGCTTGCAGTGCATG	1511
Qy	536	-----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnThr	551
Db	1517	GGGGCATCCTGTGATCTTCAGGGCTGCAC-----	154
Qy	551	rlyGlnSerGlyGlnAsnLeuProTyTrLeuValAlaTyGlnAlaIleThrValCysAlaAr	571

Db	1548	-----GGCTGGCCAAACAAGCCTCAGACGAGTATGTCGCTCA	1583
Oy	571	galaGlnAlaPro-----ProPseTrrPasp---GlnMettr	583
Db	1586	CGGCGACGACCGCTCGCTGAGTCCATCATCAGAGAGCCAAAGAGCTGGACCTGGACGAGCTG	1645
Oy	583	plyGcys 585	
Db	1646	GGCTTGC 1652	
RESULT 10			
AKO31679			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of
Direct Submission

Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

CDS

1..2984
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030486C03"
/db_xref="WGI:2394945"
/db_xref="taxon:10090"
/clone="6030486C03"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
25..2754
/note="unnamed protein product; hypothetical P-loop
containing nucleotide triphosphate hydrolyases structure
containing protein (SCOP|52540, evidence: SCOP)
putative"
/codon_start=1
/protein_id="BAC27509.1"
/db_xref="GI:26327531"
/translation="MAAGRRVGPSPPTMAPPKKRLRRRTGASQGRSDSDSE
FEIODEDARARLPGALPSPFSECVSEVPEPTBEMVAONKKKSGSGFSGWGL
SYPFKQIMKKGVPTPIORRTIPLVDGKVAMARVSGKACFLPMFELKAR
SAQTGARALILSPRELALQMTKEKFTGKLTALILGGDMEDQFALHNPDI
IATRGRLVHVAEMNLKQSEYVVEDEARLPEMGAEQLQELIGRLPGHQTVF
SATLPLKLVEPARGLTPEVPLILDNDKXINLOKTSRLVREPTKAVILYLQNV
RPOQTVFPAATKHAETVETELMGQYSCNHTISALDQTKRKINLAKFTNKSSTLI
VTDLAARGLDLPLDNIINYSFPAKGLPLHRVGVVAAAGSGTAYSLVADVPYLL
DLHLFLRSVTLARPCEPSVADAVDGVGVGVPOSVVDDESLQVAMASLDLG
LHRYANNAQOQVRSRPAESPESIKRAKELDLAELGLPLFSFCFEGEGLRLVDS
IKVYRTTTFEINASSKDPSSOMRAKORRRAKAVAFQORRQREGGPDPAPRE
LPOEEEMETVGVCTEVGVQKRRPBGSGAKRRBMETROBQDFVYVYRKPD
SERGLSVSGAGAEQVAVAVLMDGDEANMRGQQQLMDKKRFPVQSQDEBK
KRTITESGRFTISSYKDLVYKWKQKIDIRDSBERGSPNQRPGRGGRKRSQ
TSQPRASSVPAGRMRSSELKTEQILKQPRGSEALPAARGPEAAFTQPPSPGAAP
GRLPGCSLQGGQDEKDVRSQTPQWGLFVAVGVDRDYVAVHLEDPSTGAL"
2963..2968
/note="putative"
2984
/note="putative"
2984
/note="putative"

ORIGIN

Alignment Scores:

Pred. No.: 3.5 Length: 2984
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK031679 (1-2984)

QY 100 CysGlySerSerAspLeuTyrIleu-----ValThrArgHisAlaAspVal 114
|||||:.....:|||||

Db 98 TCGCAAAAGCCGAACTGGGCGCTTCCCAAGCGCCGACAGCGACTCGGATGACGGCGACT 157
QY 115 TLeProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db 158 TCGAGATCCAGCGCGAGAGATGACGCCCGCGAGAGAGCTGGGCGCTCGAGAGACCTTGC 217
QY 132 Pro-ILeSerTyrIleuLys-----GlySerSerGlyGlyProLeuLe 145
Db 218 CTTCAATTCCTACCTCAGAGTGGCGTATCATGATGTGAGCCCGACATCGGAGATGTGTGC 277
QY 145 uCyProAlaGlyHisAlaValGlyIlePheArgAlaVal----- 159
Db 278 GAGCCCGAACAAGAAAAGAGAAGAGAGTGGAGCTTCGATCGATGCGCTGATTACC 337
QY 160 -CysThr-ArgGlyValAlaGlyAlaValAspPheIleProValGlySerLeuThrT 179
Db 338 CTGTTGTCMAAGGGATATGAAAAAGGCTTACAAAGTCCGACGCCATCCAGAGAA 397
QY 179 hMetArgSerProValPheSerAspAsnSerSerProValAlaValProGlnSerTyrG 199
Db 398 CCATC-----CCGATGATCTGGATGCAAGATGTGTGGCCATG----- 438
QY 199 LmValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db 439 -----GCCCGACAGCGAGTGGCAAGCGGCTGCTCTCTCCCGA 481
QY 214 -----LysValProAlaAlaTyrAlaValGlnGlyTyrLysValLeuValLeuAsp 231
Db 482 TGTTTGAGCGGCTGAAGCAGCGAGTGCACAGAGCGGGCTCGAGCCCTCATCTTCAC 541
QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 247
Db 542 CCACCCGGAGCTGGCCCTCGACAGCATGAAGTTCATTAAGACTGAGCAAGTT-CACC 600
QY 248 -----GlyLysAspProAsnIleArgThrGlyValArgThrIleThrGlys 264
Db 601 GGCCTCAAGACAGGCTTGAATCTGGGTGGAGACAAATGAGACAGATTGGACGCCCTG 660
QY 264 eProlIethrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGly 284
Db 661 CACGAGAACCTGACATATCATTCATTCGCCACCCCTGGGCTGTGTGATGT-----GGCT 714
QY 284 LArYAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSerIleLeuGly 304
Db 715 GTGAGAGTAACTGAAGCTGACAGAGTGT-----GAGATATGTGGTG 756
QY 304 Le----- 304
Db 757 TTCATGAGACAGACAGGCTCTTTGAATGGGCTTTCGTGAGCAGCTACAGAGATGATA 816
QY 305 -----GlyThrValLeuAsp-----GlnAlaGlyThrAlaG 315
Db 817 GGGCGCTTCTGGGGGCGACACAGAGGTGTCTGTTCAGCTACATGACATGCCAACGTGTG 876
QY 315 LAlaAlaGlyLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db 877 GTGGAATTTGACCGGAGGAGGCTCACAGAGCGGTGTCAATCGCGCTGAGAGTGAAGTCC 936
QY 332 alProHisProAsnIleGlnGlyValAlaLeuSerThrThrGlyGlnIlePro----- 349
Db 937 AAGCTCAATGAGCAGCTCAAGACCTC-----CTTCCTCTCTGTG 975
QY 350 -----PheTyrGlyLysAlaAlaLeuProLeuGlnAlaIleLys-----GlyG 363
Db 976 CGCGAAGACCAAGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1035
QY 363 Lr-ArgHisLeuIlePheCysHisSerLysLysCysAspGlyLeuAlaValLeu 382
Db 1036 GACCAAGCTGT 1095
QY 383 ValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
Db 1096 ATGGGCCAGGCTGT 1146

```

Qy 403 Thrserylaspvalvalvalalathraspalaleumethrthrglypethrthrglyasp 422
Db 1147 -----GCCGCAAGTCACTGGCCAGATTCCACACACAC 1182
Qy 423 Pheaspservallileaspyasanthr---Cysvalthrghlnthrvalaspheserleu 441
Db 1183 -----AAATGTTCAACCTCATCTGTGCTGACCTGGCGCCGGGGGCTG 1227
Qy 442 AspProthrphethrthrglnthrthlethrleuproglnaspala----- 456
Db 1228 GACATCCCACTGCGGCAACAGCATACATCAAGCTTCCCTGCCAGGCAAGCTCTTC 1287
Qy 457 -----VALseragthrlnarararararararararararararararar 473
Db 1288 CTGACCCAGATGGCGCGTGGCCCGGAGCAGCGCCGAAGTGGCACAGCC----- 1335
Qy 474 Tyrarphelvalalaprogllyglu-Argproserglymeapheaspserservalley 493
Db 1336 TATCTTGTGGTGGCCGAGACAGAGTCCCTAC----- 1369
Qy 493 sglucytyraspalaglycysalatptrygluleuthrproalagluthrthralar 513
Db 1370 -----TGCTTACCTACACCTGCTGGCGCTGTCTGTACCTGGCCCTCTTGTGAG 1425
Qy 513 g---Leuaralatyrmethanthrproglyleuprovalcysglnasphisleugluph 532
Db 1426 GAGCCTTCAGTGGCAGATGGCGTGGCAGGG-----ACCGAGTGG 1464
Qy 532 eTropglugly----- 535
Db 1465 CTGGTCCGCGTGGCCCAAGTGTAGTGTATGATGATGAGACAGACAGCTGCAGATGCCATG 1524
Qy 536 -----VALPhehrthrglyleuthrthrsileaspalahispheluseerghnth 551
Db 1525 GGGGCGATCCCTGATCTTCAGGGCGCTGCAC----- 1555
Qy 551 rlygslnserglyluamleuprotyleuvalalatyrglnalathrvalcysalaar 571
Db 1556 -----GCCTGGCCAAACACGCTCAGCAGCACTATGTGGCTCA 1593
Qy 571 gAlaglnalapro-----ProboSerTTPasp---GlnMetTr 583
Db 1594 CGGCGACGGCGCTCGCTGATGATCATCAAGAGCCAGAGAGCTGAGCTGGCAGAGCTG 1653
Qy 583 pLyScys 585
Db 1654 GCGCTTGC 1660

```

```

RESULT 11
AK028274 2986 bp mRNA linear HTC 18-SEP-2003
LOCUS Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
DEFINITION enriched library, clone:3222401E18 product:hypothetical P-loop
containing nucleotide triphosphate hydrolases structure containing
protein, full insert sequence.
ACCESSION AK028274
VERSION AK028274.1 GI:26080791
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

```

```

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komuro, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichipillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 2986)
Adachi, T., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshitake Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .2986
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:3222401E18"
/db_xref="MG1:2350122"
/db_xref="taxon:10090"
/clone="3222401E18"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="14, 17 days embryo"
21. .2751
/note="hypothetical P-loop containing nucleotide
triphosphate hydrolases structure containing protein
(SCOP|52540, evidence: SCOP)
misc_feature

```

polyA_signal putative"
2960..2965
/note="putative"
polyA_site 2986
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 3.5 Length: 2986
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
Gaps: 29

US-09-930-591-2 (1-686) x AK028274 (1-2986)

QY 100 CysglYserSerApleuTYrleu-----ValTHrArgHISAlaAspVal 114
Db 94 TGCAGAAACGCCGAACTGGGCTTCCCAAGCCCGCAGCAGCACTCGGATGACGGCAGT 153
QY 115 IleProValArgAArg-----GlyAspGlyArgGlySerLeuSerProArg 131
Db 154 TCGAGATCCAGCGAGGATGACGCCCGGCGAGAGCTGGCTTGGCAGAGCTTGC 213
QY 132 Pro-IleSerTYrleuYs-----GlySerSerGlyGlyProleuLe 145
Db 214 CCTGATTTCTACCTCAGAGCGCGATGAGATGAGACCCGACACTCGGAGATGCGC 273
QY 145 uCyProAlaGlyHISAlaValGlyLePheArgAlaAlaVal----- 159
Db 274 GAGCCCGAAGCAAGAAAGAAAGAAAGTCTGAGGCTTCAGTCCATGGGCTGAGTTAC 333
QY 160 -CysThr-ArgGlyValAlaAlaValAlaPheIleProValGluSerLeuGluThr 179
Db 334 CTGGTTCAAGGGATCTGAAAGGGCTACAGAGTCCGACGCCCATCCAGAGAGAA 393
QY 179 hrMetArgSerProValPheSerAspAsenSerSerProProAlaValProGlnSerTYrG 199
Db 394 CCATC-----CCGTGATCTTGATGGCAAGATGTGTGCGCATG----- 434
QY 199 InValAlaHISleuHISAlaProThrGlySerGlyYsSerThr----- 213
Db 435 -----GCCCGACAGGCGAGTGGCAGACGCGCTGCTTCTCTCCCGA 477
QY 214 -----LysValProAlaAlaAlaAlaGlnGlyTYrLysValLeuValLeuAsp 231
Db 478 TGTTTGAGCGGCTAAGGACAGAGTGCACAGACGCGGCTCGAGCCCTCATCTCTAC 537
QY 231 roser-----ValAlaAlaThrMetGlyPheGlyAlaTYrMetSerLysAlaHISG 248
Db 538 CCACCCCGGAGCTGGCCCTGCAGACCATGAACTTACATTAAGACTAGGCAAGTT-CAAC 596
QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlyS 264
Db 597 GGCCTCAAGACTGCTTATCTGGTGGAGCAAAATGAAAGCAAGTTTGGACGCTTG 656
QY 264 expProIleThrTYrSerThrTYrGlyYsPheLeuAlaAspGlyGlyYsSerGlyGly 284
Db 657 CACAGAAACCTGACATATATGCAACCCCTGGGCTGTGGTCAAGT-----GGCT 710
QY 284 laTYrAspIleIleIleCysAspGluCysHISerThrAspAlaThrSerIleLeuGlyI 304
Db 711 GTGAGATGAACCTTGAAGCTGCAGATGT-----GAGATATGTGTG 752
QY 304 le----- 304
Db 753 TTGATGAGACAGACAGGCTTTGAATGGCTTTGTGAGAGACTACAGAGATCAT 812
QY 305 -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db 813 GCGCGCTTCTCTGGGGGCCACACAGCGGTGCTGTTCTCAGTACACCTGCCAAGCTGCTG 872

QY 315 lYAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
Db 873 GTGGAATTTGCACGGGACGCGCTTCACAGAGCCCGCTCTCATCCGCTTGAAGTACTCC 932
QY 332 alProHISProAsnIleGluGluValAlaLeuSerThrThrGlyGluIlePro----- 349
Db 933 AAGCTCAATGACGAGCTCAAGACTC-----CTTCTCTCTGTG 971
QY 350 -----PheTYrGlyValAlaIleProLeuGluAlaIleLys-----GlyG 363
Db 972 CGCAGAACACCAAGGCTGCTGCTCTCTACTCTGCGAGATGCTGGGCCCCAG 1031
QY 363 ly-ArgHISleuIlePheCysHISerLysLysLysCysAspGluLeuAlaAlaLysLeu 382
Db 1032 GACCAAGCTGGGTGTTGTGAGCAAAAGACCATGCGAGTACTCAGAGATGCTG 1091
QY 383 ValAlaLeuGlyValAlaAlaValAlaIleTYrTYrArgGlyLeuAspValSerValIlePro 402
Db 1092 ATGGGCGAGGTGTGAGATTTGGCCACATCTATAGTCTTGAACACAGAG----- 1142
QY 403 ThrSerGlyAspValValAlaIleThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
Db 1143 -----GCCCGAAGTCAACTTGGCCAACTTGCACACAC 1178
QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db 1179 -----AATGTTCCACCTCATGCTGACTGACCTGCGCGCGGCGCTG 1223
QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db 1224 GACATCCCACTGTCGACCAAGTCATCACTCAAGCTTCCCTGCAAGGCGCAAGCTTTC 1283
QY 457 -----ValSerArgThrGlnArgArgGlyTYrArgThrGlyArgGlyLysProGlyIle 473
Db 1284 CTGCACGAGATGGCGCTGTGGCCCGGCGAGCGGCGGAGTGGCACAGCC----- 1331
QY 474 TYrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493
Db 1332 TATCTTGTGTGGCCGACAGAGGTCCCTAC----- 1365
QY 493 sGluCysTYrAspAlaGlyCysAlaATPTYrGluLeuThrProAlaGluThrValAr 513
Db 1366 -----TGCCTGACCTACACCTGTTCTCTGGCGGCTCTGCACTCGGCCGCTTGTGAG 1421
QY 513 g---LeuArgAlaLeuThrAsnThrProGlyLeuProValCysGlnAspHISleuGluPh 532
Db 1422 GAGCTTCAAGTGGCAGATGGGTTGGCAGG-----ACGAGATG 1460
QY 532 eTPGluGly----- 535
Db 1461 CTGGGTGGCGGCCCGCAGAGTGTAGTATGATGAGACAGACGCTGCAGATGCCATG 1520
QY 536 -----ValPheThrGlyLeuThrHISIleAspAlaHISPheLeuSerGlnTh 551
Db 1521 GGGGATCCCTGATCTTCAAGGCGCTGCAC----- 1551
QY 551 rlyGlnSerGlyLysAsnLeuProTYrLeuValAlaTYrGlnAlaThrValCysAlaAr 571
Db 1552 -----GCTGGCCAAACACCTCAAGCAGATATGTGGCTCA 1589
QY 571 gAlaGlnAlaPro-----ProProSerTrpAsp---GlnMetTr 583
Db 1590 CGGCGACGCGCTGCTGATGCATCAAGAGCCAAAGAGCTGCACTGGCAGACTG 1649
QY 583 pLysCys 585
Db 1650 GGCCTTC 1656

RESULT 12
LOCUS BC043699 3956 bp mRNA linear HTC 13-JAN-2003
DEFINITION Mus musculus, similar to RIKEN cDNA 2410015A15 gene, clone
IMAGE:5344158, mRNA.

QY 442 AspProThrPheThrIleGluThrIlePheLeuProGlnAspAla----- 456
 DB 1210 GACATCCACATGCTGGACAGCATATCAATCAAGCTTCCTCCGCAAGGCAAGCTCTTC 1269
 QY 457 -----ValSerArgThrGlnArgArgIleArgThrGlyArgGlyLeuProGlyIle 473
 DB 1270 CTGACCCAGATGGGCGTGTGGCCGACAGCCGCAATGGCACAGCC----- 1317
 QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493
 DB 1318 TATCTTTGGTGGCCGACAGAGGTCCTCCAC----- 1351
 QY 493 sgluCySerYzAspAlaGlyCyValatPtyrGlyLeuThrProAlaGluThrThrValAr 513
 DB 1352 -----TGCTTGACCTACACACTGTCTCTGGCGGCTGTGCACCTGGCCGCTTGTAG 1407
 QY 513 g---LeuArgAlaTyrMetAsnThrProGlyLeuProValCySglAspHisLeuGluPh 532
 DB 1408 GAGCCTTCAGTGGGAGATGCGGTTGGCAGGG-----ACGAGTGG 1446
 QY 532 eTPrglGly----- 535
 DB 1447 CTGGGTGCGGTCGCCACAGTGTAGTGAATGAGACAGACGCTGCAGACTGCCATG 1506
 QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
 DB 1507 GGGGCAATCCCTGATCTTCAGGGCCTGCACC----- 1537
 QY 551 rlyGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCyValAr 571
 DB 1538 -----GCGTGGCCAAACGCTCAGCAGCAAGTATGATGGCTCA 1575
 QY 571 galGlnAlaPro-----ProProSerTyrPAP---GlnMetTr 583
 DB 1576 CGGCCAGGCGCTCGCTGAGTCCATCAAGAGCCAGGAGCTGCAGCTGGCAGAGCTG 1635
 QY 583 pLyseCy 585
 DB 1636 GGCCTTGC 1642

RESULT 13
 AY404177 1697 bp DNA linear GSS 15-DEC-2003
 LOCUS Mus musculus DX34 gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY404177.1 GI:39760154
 VERSION AY404177.1
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1697)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Cavello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Interfering nonneutral evolution from human-chimp-mouse orthologous
 gene tritos
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1697)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Cavello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 COMMENT Rockville, MD 20850, USA
 FEATURES This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 location/Qualifiers

source 1. 1697
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1.1697
 /gene="DX34"
 /locus_tag="HGM1803"

gene

ORIGIN

Alignment Scores:

Pred. No.:	1.44	Length:	1697
Score:	128.50	Matches:	134
Percent Similarity:	35.56%	Conservative:	68
Best Local Similarity:	23.59%	Mismatches:	221
Query Match:	3.55%	Indels:	148
DB:	29	Gaps:	30

US-09-930-591-2 (1-686) x AY404177 (1-1697)

QY 52 ValCySerPthrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLyGlyPro 71
 DB 116 GTCTCTGGAAGATGATTTCTCCGCG--ATGAGATTTATTC-----GTC 160
 QY 72 ValIleGln-----MetTyrThrAsnValAspGlnAspLeuVal 84
 DB 161 GTGGTTCAGAGAGAGTGCAGAAAGTTCTGGAGCTTTCTTGAAACGCTGCAGAGATTCAGC 220
 QY 85 GLyTyrProAlaProGlnGlnGlyAlaArgSerLeuThrProCySerTyrCySglYserSerAsp 104
 DB 221 ACCTCAAGACCTCCCAAGAGAAGAGAGAACCTGTCATGCCCAACATGGCCTGCTG 280
 QY 105 Leu-----TyrLeu---ValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 281 CTTTAGCTGACCTACCTCTCACTTACAGCCACGCT----- 316
 QY 121 GLyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLyGlySerSer 140
 DB 317 ---ACGCAATCAACCTCTCATCTTCAGCCAGACACTCGGGGTCGATGGCCGCGCA 373
 QY 141 GLyGly-----ProLeuLeuCyProAlaGlyHisAlaValGlyIlePheArgAlaAla 158
 DB 374 GGGGGCTGCCCCCGAGAGAGTGTCCAGTTC--GCGGTGCAATTCCTTCACTACCTAG 430
 QY 159 ValCySerThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeu---Glu 177
 DB 431 ACTTCCAGCAGA-----ACAGGCTTTGGAGGCTCCGCAACTGACAGCTGAGGAG 484
 QY 178 Thr-ThrMetArgSerProValPheSerAspAsnSerSerProAlaValProGln-S 197
 DB 485 CGGCACTTCCCATGGCCAGATGGGAAACGCAAT-----CTGCAGACTCTCAAG 535
 QY 197 eTyrGlnValAlaHisLeuHisIleAlaProThrGlySerGlyLySerThrLyValProAl 217
 DB 536 AACACAGTGGTGGTGGTGGTGGCCGGGACAGAGCTGTGGCAAGTCTACCTACAGGTGCC 595
 QY 217 lalaTyrAlaAlaGlnGlyTyrLyS---ValLeuValLeuAsnPro----- 231
 DB 596 AGTACTTCTGGCGCGCGGCTTCAAGTCACTGGCATGCACTCAAGCCCGAGAAATGCGCT 655
 QY 232 -----SerValAlaAlaAlaThrMetGlyPheGlyAlaTyrMetSerLyAlaHisGlyIleA 250
 DB 656 GCATCTCACTAGGCGCAAGCGGCTTGGAGAGCCGATGATAGGCTCCAGAGTTG 715
 QY 250 sProAlaHisGlnGlyValAlaArgThrIleThrGlySerProIleThrTyrSerT 270
 DB 716 GCTACAGATCCGCTTT--GAAGGCACTGGTGGGAGGCGGCAACATGATGTGTCTGTA 772
 QY 270 hTyrGlyLyPheLeu-----AlaAspGlyGlyCySerSerGlyGlyAlaTyrAspIle 288
 DB 773 CGGTGGGCTGCTCTTACGCGCAATCCAGGCTGAGAGCCAGCTGCCCCAATGACAGTCC 832
 QY 288 leIleCyAspGlyCyHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrVal 308

Db 385 GGTGTGAAGCTTGGAGAGAGTGGCTACAGATCCGGTTTGAGACTGCACCTCAGAG 444
 Qy 258 ArgThrIlethrThrGlySerProIleThrTySerThrTyGly-----Lys 273
 Db 445 CGAAGTGT-----CTCCGCTACAGACAGATGAGATGCTACTCCGAGAG 489
 Qy 274 PheLeuAlaAspGlyGlyCysSerGlyAlaTyraAspIleIleCysAspGlyCys 293
 Db 490 TTCCTCTGTGAGCTGACCTTGCA-----AGTTACAGTGTGTGATGATGATGAGACT 543
 Qy 294 HisSerThrAspAlaThrSerIleuGlyIleGlyThrValLeuAspGlnAlaIuThr 313
 Db 544 CACGAGCGGACCTTGACACAGACATCTCTTTGATTGATTCAGAGACGTCGTAGATTC 603
 Qy 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 Db 604 CGACCTGAGCTCAAGTCTGTGTGCTTCAGCCACA 639
 RESULT 15
 CD240900 738 bp mRNA linear EST 22-MAY-2003
 LOCUS AGENCOURT 14124904 NIH MGC 177 Mus musculus cDNA clone
 DEFINITION IMAGE:30380772 5', mRNA sequence.
 ACCESSION CD240900
 VERSION CD240900.1 GI:31001364
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE NIH-MGC
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Michael Brownstein Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: NDCM173 row: f column: 13
 High quality sequence stop: 482.
 Location/Qualifiers
 1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30380772"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH_MGC_17"
 /note="Organ: liver; Vector: pDNR-LIB; Site 1: SfiI
 (ggcgatcgcc); Site 2: SfiI (ggcgcccgcc); cDNA made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGACAGTGCATTCAGGCGGG-3' and
 5'-ATTCTAGAGCGGAGCGGCGGACATG-dT(30)NN-3'. Full-length
 enriched library was constructed using the Clontech
 Creator SMART kit and size-selected to contain the 0.5 kb
 size fraction. Library created in the laboratory of M.
 Brownstein (NIH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 0.414 Length: 738
 Score: 127.50 Matches: 45
 Percent Similarity: 44.19% Conservative: 31
 Best Local Similarity: 26.16% Mismatches: 69
 Query Match: 3.52% Indels: 27
 DB: 14 Gaps: 7

US-09-930-591-2 (1-686) x CD240900 (1-738)
 Qy 174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerProPro 192
 Db 62 GAATCTATCCAGGCTGTGCGCGGACGCTGCCCTGTTCCCTCCAGAGAGCTTCTG 121
 Qy 193 AlaValProGlnSerTyGlnValAlaHisIleuHisAlaProThrGlySerGlyLysSer 212
 Db 122 GCCGCCATTGCCAACACATCAGAGTCTCATCATCGAAGGCGAGACTGCTCGGGAAGACC 181
 Qy 213 ThrIleValPro-----AlaAlaTyraAlaAlaGlnGlyTyTyIleValLeu 227
 Db 182 ACACAGATCCACAGATCCTTTGAGAGAGGTTACCAAGAAAGGCGATGAGATTCCT 241
 Qy 228 ValLeuAspProSerValAlaAlaThrMetGlyPheGlyAlaTyMetSerIleAlaHis 247
 Db 242 TGCACCCAGGCCCGAGAGATGCGGCTATGATGTGGACGCGCAGATGCGCCGAGATG 301
 Qy 248 Gly-----IleAspProAsnIleArg-----ThrGlyVal 257
 Db 302 GGTGTGAAGCTTGGGAGAGAGTGGCTACAGATCCGGTTTGAGACTGCACCTCAGAG 361
 Qy 258 ArgThrIlethrThrGlySerProIleThrTySerThrTyGly-----Lys 273
 Db 362 CGAAGTGT-----CTCCGCTACATGACAGATGAGATGCTACTCCGAGAG 406
 Qy 274 PheLeuAlaAspGlyGlyCysSerGlyValTyraAspIleIleCysAspGlyCys 293
 Db 407 TTCCTCTGTGAGCTGACCTTGCA-----AGTTACAGTGTGTGATGATGATGAGACT 460
 Qy 294 HisSerThrAspAlaThrSerIleuGlyIleGlyThrValLeuAspGlnAlaIuThr 313
 Db 461 CACGAGCGGACCTTGACACAGACATCTCTTGTGATTGATTCAGAGACGTCGTAGATTC 520
 Qy 314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
 Db 521 CGACCTGAGCTCAAGTCTGTGTGCTTCAGCCACA 556

Search completed: September 17, 2004, 12:08:02
 Job time : 5807 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 06:45:20 ; Search time 7206 Seconds
(without alignments)
4126.189 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPTAYAOCTRGHLCIT.....PAIPDRVLYRFEDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgm2_1/USFTO.epool/US0930591/runat_13092004_164952_688/app_query.fasta_1.839
-DB=genbml -QWrt=faetap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=ppc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US0930591 @CGN 1.1 4390 @runat_13092004_164952_688 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genbml:*
1: gb_ba:*
2: gb_ncg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3618	100.0	2061	6 AX441176	AX441176 Sequence
2	3618	100.0	2061	6 AX467113	AX467113 Sequence
3	3574	98.8	5360	6 I06434	I06434 Sequence 48
4	3574	98.8	5360	6 I09328	I09328 Sequence 8
5	3574	98.8	6785	6 I06440	I06440 Sequence 54
6	3574	98.8	6785	6 I09329	I09329 Sequence 10
7	3574	98.8	7310	6 AR118696	AR118696 Sequence
8	3574	98.8	7310	6 I09331	I09331 Sequence 15
9	3574	98.8	7310	14 HPCPOLYP	M32084 Hepatitis C
10	3574	98.8	9185	6 BD08294	BD08294 Sequence 1
11	3574	98.8	9185	6 BD091382	BD091382 HCV culti
12	3574	98.8	9379	6 AR166930	AR166930 Sequence
13	3574	98.8	9379	6 AR301300	AR301300 Sequence
14	3574	98.8	9401	6 AR176483	AR176483 Sequence
15	3574	98.8	9401	6 E66593	E66593 Hepatitis C
16	3574	98.8	9401	6 I71894	I71894 Sequence 9
17	3574	98.8	9401	6 I81885	I81885 Sequence 9
18	3574	98.8	9401	6 BD080334	BD080334 Hepatitis
19	3574	98.8	9609	12 HPCPOLYP	M62321 Hepatitis C
20	3574	98.8	9609	12 AF387805	AF387805 Synthetic
21	3574	98.8	9646	12 AF387806	AF387806 Synthetic
22	3574	98.8	9693	12 AF387807	AF387807 Synthetic
23	3571	98.7	6785	6 AR118692	AR118692 Sequence
24	3570	98.7	8316	6 AR118703	AR118703 Sequence
25	3570	98.7	9185	6 AR118722	AR118722 Sequence
26	3570	98.7	9185	6 AR118723	AR118723 Sequence
27	3570	98.7	9618	14 AF271632	AF271632 Hepatitis
28	3568	98.6	6299	6 AX164584	AX164584 Sequence
29	3565	98.5	2058	6 AR404933	AR404933 Sequence
30	3565	98.5	2058	6 AR408362	AR408362 Sequence
31	3565	98.5	2058	6 AX395309	AX395309 Sequence
32	3565	98.5	2058	6 AX454818	AX454818 Sequence
33	3565	98.5	8987	6 AR118728	AR118728 Sequence
34	3565	98.5	9609	12 AF387808	AF387808 Synthetic
35	3555	98.3	5360	6 AR118686	AR118686 Sequence
36	3551	98.1	9646	6 AR110828	AR110828 Sequence
37	3551	98.1	9646	6 BD069982	BD069982 Functiona
38	3551	98.1	12980	6 AF009606	AF009606 Hepatitis
39	3551	98.1	12980	6 AR110831	AR110831 Sequence
40	3551	98.1	12980	6 BD069985	BD069985 Functiona
41	3550	98.1	9395	14 AF511950	AF511950 Hepatitis
42	3550	98.1	9502	6 E08263	E08263 gRNA of Hep
43	3550	98.1	9502	6 E08264	E08264 CDNA of Hep
44	3550	98.1	9502	14 HPCUCJ1	D10749 Hepatitis C
45	3547	98.0	9610	14 HEC278830	AJ278830 Hepatitis

RESULT 1

ALIGNMENTS

AX441176
LOCUS AX441176 2061 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 16 from Patent WO0213855.
ACCESSION AX441176
VERSION AX441176.1 GI:21665758
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Sallberg, M. and Hultgren, C.
TITLE Vaccines containing ribavirin and methods of use thereof
JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;
TRIPEP AB (SE)
FEATURES
source location/Qualifiers
1..2061
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Hepatitis C virus NS3/4A coding region"
ORIGIN
Alignment Scores:
Pred. No.: 1,41e-203 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-930-591-2 (1-686) x AX441176 (1-2061)
QY 1 MetAlapProIleThrAlaTYrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
Db 1 ATGGCGCTTATCAGCGCTATGCCAGACAGAGGGGCTTTGGATGATATAC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGlyValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGCAAAAACAGGTGGAGGTGAGTTGATCGATCGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysGlnLeuGlyValCysTrpThrValTYrHisGlyAla 60
Db 121 GCCAGACTTCTTGGCAACTGATTAACGGGGTGTGACTGCTACATGAGACC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTYrThrAsnValAsp 80
Db 181 GGAACAGAGACCATTTGGCTACCTAAGGCTCTGTATCCAGATGACCAAGTGTGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTCGTAGGCTGGCCGCCGCCAGAGTCCGCTCATTAACACCATGACCTTGC 300
QY 101 GlySerSerAspLeuTYrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCTCTGGACCTTACCTGATCAGAGGACGCCGATGTCATTTCTGTGCGCGAGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTYrLeuLysGlySerSer 140
Db 361 GGTATGCGAAGGGGAGGCTGCTTCCGCCGCCCTATCTTAAAGGCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAGGCGCTCTGCTGTGCGCGCGAGCATCCGTAAGGCATATTCAAGCGCGGTATGC 480
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGlnSerLeuGlnThrThMet 180
Db 481 ACCGTGGAGTGGCTTAAGCGGTGACTTCACTCCCGTAAGAGCTTGAAGACAACATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTYrGlnVal 200
Db 541 AGGTCCCGGTGTCTTCAAGACACTCTCCCAACAGAGAGTGGCCCAAGCTCAACAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTYrAla 220

Db 601 GCCACCTGCATGCTCCACCGCGAGCGGTAAAGACACCAAGTCCCGGCCCATACGA 660
QY 221 AlaGlnGlyTYrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTTACAAAGGTGTGGTGTCAACCCCTCGTTGTCTGCACAAATGGGCTTTGGT 720
QY 241 AlaTYrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAAGGCCCATGGATGTATCTTAACATCAGAGCTGGGGTGAAGCAATT 780
QY 261 ThrThrGlySerProIleThrTYrSerThrTYrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACTACTGGGAGCGCCGATCAAGATTCACCTACCGGCAAGTTCTTGGCAGCGGGGTGT 840
QY 281 SerGlyValAlaTYrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGGTCTTATATACATTAATTTGTGAGAGTGCACATCCAGATGCACATC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCATTTGGCAGCTGCTCTTGACCAAGCAGAGACCGCGGGCGAGCTGACTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
Db 961 CTGCGACCGCTACCCCTCGGGCTCGTCACTGTGCCCCATCTTAATCGAGAGTT 1020
QY 341 AlaLeuSerThrArgGlyGluIleProPheTYrGlyLysAlaIleProLeuGlnAlaIle 360
Db 1021 GCTCTGTCACTACCGGAGAGATCCCTTTATGGCAAGCTATTCCTTGAAGCAATT 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
Db 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAAAGAAAGTGCACAGCTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTYrTYrArgGlyLeuAspValSerVal 400
Db 1141 AAACGTGTGCGTGGAGCGTCAGATGCGGTACTTACCGGCTTATGTGTCTGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACCAAGTGTGACGTTCGTCGTGGCAATGACGCCCTCATGACCGGCTTACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrGlnThrValAspPheSer 440
Db 1261 GCGGACTTCATTCGGTGTGATGACTCAACACGTGTGTACCCAGACAGTCACTTCAG 1320
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTGACCTTACCTTACCATTTGAGACATACCGCTTCCCAAGATGCTGTCTCCGTACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgLysProGlyIleTYrArgPheValAlaProGly 480
Db 1381 CAAGCTCGGGGTAGCATGGCAGAGGGAAGCAGGATCATCAGATTTGTGGACCGGGG 1440
QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlnCysTYrAspAlaGlyCys 500
Db 1441 GAGCGTCTTCTGCAATGTTGACTGCTGTCTGTGCGAGTGTGATGACGGGGTGT 1500
QY 501 AlaTrpTYrGlnLeuThrProAlaGluThrThrValArgLeuArgAlaTYrMetAsnThr 520
Db 1501 GCTTGGTATAGCTTACCGCCCGGAGACCAACAGTTAAGCTACGACATACATGAACAC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheTrpGlnGlyValPheThrGlyLeu 540
Db 1561 CCGGAGCTTCCCGTGTGCCAAGACACTTGAATTTTGGAGGGCGCTTTTACGGGTCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGlnLeuProTYr 560
Db 1621 ACCCACTAAGCCCACTTCTATCCCAAGAGAGTGGGAAACCTTCCCTAT 1680
QY 561 LeuValAlaTYrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580

Db 1681 CTGGTAGCGTACCAAGCCACCGGTGTGGCTAGAGCTCAAGCCCTCCCGGTGGGAC 1740
 Qy 581 GIMeTTrpLysCysLeuIleArgLeuLeuProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGGAAAGTCTTGATCCGCTCAAGCCCACTTCATGGGCCCAACCTCTCTTA 1800
 Qy 601 TTAArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 1801 TATAGACTGGGCGGTGTCCAGAAATGAAGTCACTTGACGACCCAGTCAACAATATATC 1860
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValIleValGlyGly 640
 Db 1861 ATGACATGTATGTGCGGTGACCTGGAGGTGCTCAAGATACCTGGGTCTGTGGCGGC 1920
 Qy 641 ValLeuAlaAlaLeuAlaValTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 1921 GTTCTGCTGCTTTGGCGCGGTATGGCTATCCACAGGCTGCGGTGATAGTATAGG 1980
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGAAAGCCGGCAATCATACCGACAGGGAAGTCTCTACCGGGAATTTC 2040
 Qy 681 AspGluMetGluGluCys 686
 Db 2041 GATGAAATGGAAAGTGC 2058
 RESULT 2
 AX467113
 LOCUS AX467113 2061 bp DNA linear PAT 17-JUL-2002
 DEFINITION Sequence 1 from Patent WO0214362.
 ACCESSION AX467113
 VERSION AX467113.1 GI:21900425
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Sallberg, M.
 TITLE A hepatitis C virus non-structural ns3/4a fusion gene
 JOURNAL Patent: WO 0214362-A 1 21-FEB-2002;
 TRIPEP AB (SE)
 FEATURES
 source 1..2061
 Location/Qualifiers
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Hepatitis C virus NS3/4A coding region"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,41e-203 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-530-591-2 (1-686) x AX467113 (1-2061)
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 1 ATGGCGCTATCAAGCGCTATGCCAGACAGCAAGGGGCTTTGGATGATATATACCC 60
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 61 AGCTTGAACCGGCGCGGACAAACACAGGTGAGGGGTGAGTTCAAGATCGTCAACTGCT 120
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 121 GCCAGACTTCTTGGCACTGCACTTACCGGGGTGTGTGACTGTCTACCAATGAGACC 180
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

Db 181 GGAACAGAGCAATTGCGTACCTAAGGTCCTGTTATTCAGATGTACCAATGTGAC 240
 Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 241 CAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCGCTCATTAACCATGCACTTGG 300
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 301 GGCTCTCGGACTTAACTTACCTGTGCAGAGGACGCGGATGTCACTTGTGGCCCGCAGC 360
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 361 GGTAATGGCAGGGGACGCTGCTTGGCCCGCTTATCTTATCTTGAAGGCTCTTGG 420
 Qy 141 GlyGlyProLeuLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 421 GGAGGCTCTGCTGTGTCGCGGAGACATGCGGTAGGCAATTCAGAGCCGGGTATGC 480
 Qy 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 481 ACCCGTAGAGTGTGCTAAGGCGGTGACTTCACTCCGCTAGAGCTTGAAGACACATG 540
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 541 AGGTCCCGGTGTCTCAGCAACTCTCCCAACAGAGTGTCCAGAGCTACCAAGTG 600
 Qy 201 AlaHisLeuHisAlaIleProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 601 GCCCACTCGATCTCTCCACCGGACGGTGAAGACCAAGTCCCGCGCATACGCA 660
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 661 GCTCAGGGCTACAAAGGTGTGGTGTCAACCCCTCGGTGTCTCAACAATGGGCTTTGGT 720
 Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaGlyThrIle 260
 Db 721 GCTTACATGTCCAAAGGCCCATGGGATGTATCTTAAATCAGACTGGGTGAGACAAATT 780
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 781 ACTACTGGAGCGCCGATCAAGTATTCACCTTACGCGCAAGTCTTGTCCAGCGCGGTGT 840
 Qy 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 841 TCAGGGGGGTCTTATGACATTAATATTGTGAGAGAGGCACTCCACGATGCAATCTCC 900
 Qy 301 IleLeuGlyTyrIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgGluThrVal 320
 Db 901 ATCTTGGCAATGGCACTGTCTTGTACCAAGCAGACACCGCGGGCGAGACTGCTGTG 960
 Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 961 CTCGCAACGCTACCTCCGAGCTCGGTGACATGTGCTGCTTCAATCAGAGAGTT 1020
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 1021 GCTCTGCCACTACCGGAGAGATCCCTTTATGGCAAGGCTATTCCTTGAAGCAATT 1080
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 1081 AAGGGGGGAGACATCTCTCTGCACTCACTCAAAAGAGAGTCAACGAGCTCGCGCA 1140
 Qy 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 1141 AAATGTGTGCGTGTGGCGCAATGCGGTGACTTACTACCGCGCTTGAATGTGCTGTC 1200
 Qy 401 IleProThrSerGlyAspValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACCAAGTGTGATCGTGTGTGTGCGCAATGAGCGCTTCAAGACCGGCTTATAC 1260
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 1261 GCGCACTTCATTCGATGTAGTACGCAACACGTGTGTACCCAGACGTGCACTTCAGC 1320

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACCTTACCATTTAGACATACACGCTTCCAGAGATGCTGCTCCGACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyValProGlyIleTyrArgPheValAlaProGly 480
 Db 1381 CAGGTCGGGGTACGATGCGACAGAGGAGACGACATCTACAGATTGTGGACCGGG 1440
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyGlyCysTyrAspAlaGlyCys 500
 Db 1441 GAGGCTCTTCTGACATGTTTGACTCGCTGCTCTGACAGTGTATGACGCGGGTGT 1500
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 1501 GCTTGGATGAGCTTACGCGCGGACCAAGACAGTACAGTACAGATACATGACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheTrpGlyLeu 540
 Db 1561 CCGGACCTTCCCGTGTGACAGACCATCTTGAAATTTGGAGGGCGCTTTTACGGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGluAsnLeuProTyr 560
 Db 1621 ACCACATAGACGCGCCACTTCTATCCAGCAAGACAGAGTGGGAAAACTTCCCTAT 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 1681 CTGTAGGCTACCAAGCCACCTGTGCTAGACTCAAGCCCTCCCGCTGTGGAC 1740
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGAAGTGTGATTCGCTCAAGCCACCTCCATGGGGCAACCTCTGCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrValTyrIle 620
 Db 1801 TATAGACGGGCGCTGTCAGAAATGATCACCTGACGACCCAGTCACCAAGTATATC 1860
 QY 621 MetTrpCysMetSerAlaAspLeuGluValAlaThrSerThrTrpValLeuValGlyGly 640
 Db 1861 ATGACATGTATGTCCGCTGACCTGAGGTCTGACAGATACCTGGGTCTGTTGGCGGC 1920
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValAlaGlyArg 660
 Db 1921 GTTCTGGCTGTTGGCGCGATATGCTATCCACAGGCTCGTGTGATGTAGTGG 1980
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTGTCCGAAAGCGGCAATCAACCCGACAGAGAGTCTTACCGGAGTTC 2040
 QY 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGAGTGC 2058
 RESULT 3
 LOCUS 106434 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 48 from Patent EP 0318216.
 ACCESSION 106434
 VERSION 106434.1 GI:590311
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 5360)
 AUTHORS Houghton M., Choo O.-L. and Kuo G.
 TITLE Nanby diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989;
 FEATURES
 1..5360 Location/Qualifiers
 /organism="unknown"
 /mol_type="unassigned DNA"

Alignment Scores:
 Pred. No.: 1.74e-200 Length: 5360
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x 106434 (1-5360)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 930 CTTGCGCCCATACAGCGCGTACCGCCAGACAGACAGGAGGCTCTCTGGGTGATATAC 989
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 990 AGCTTACTGCGCGGACCAAAACCAAGTGAAGGTGAGGTCCAGATTGTGCACTGCT 1049
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
 Db 1050 GCCCAACCTTCTGCGCAACGTGACATCAATGGGGTGTGTGACTGTACACGAGGCTC 1109
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1110 GGAACGAGACCATTCGCTACCCAGAGGTCTCTGATCCAGATGATATACATGTAGAC 1169
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1170 CAAAGACTTGTGGGCTGCGCCGCTCGCAAGTAGAGCTGCTCATTAACACCTTGACCTTGC 1229
 QY 101 GlySerSerAspLeuTyrIleValIleThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1230 GGCTCTCGACCTTACCTGATCGACAGAGCAGCGGAGTCAATTCCTCGTGGCGGGCG 1289
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleLysGlySerSer 140
 Db 1290 GGTGATAGACGGGGGACGCTGCTGCTGCCCGCCATTCTCACTTGAAGGCTTCCTCG 1349
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1350 GGGGGTCCGCTGTGTGGCCCGGGGGGACGCGGTGGGCAATTTAGGGCCGGGTGTGC 1409
 QY 161 ThrArgGlyValAlaValAspPheIleProValGlySerLeuGluThrThrMet 180
 Db 1410 ACCGTGAGTGTAGGGGGTGGAGCTTATTCCTGTGAGAGACCTTAGACACACAG 1469
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1470 AGGTCCCGGTGTTCACGATTACTCCTCCACAGTAGTGCCGACAGGCTTCCAGGGT 1529
 QY 201 AlaHisLeuHisAlaProThrArgLysSerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1530 GCTCACCTCCATGCTCCACAGGACGCGCAAAAGACCAAGGTCCGGCTGATATGCA 1589
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1590 GCTCAGGGCTATTAAGGTGCTAGTACCAACCTCTGTTGTGTCACACCTGGCTTGGT 1649
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValaGlyThrIle 260
 Db 1650 GCTTACATGTCAAGGCTCATGGAGATCGATCTTACATCGACCGGGGTGAACAATA 1709
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1710 ACCACTGGACGCCCATACAGTACTCCACTACGAGGAGTCTTGTCCGACGCGGGGTGC 1769
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaTrpSer 300
 Db 1770 TCGGGGGGCGCTTATGACATATATTTTGAAGAGTGCACATCCAGGATGCACTCC 1829
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValaGlyLeuThrVal 320
 Db 1830 ATCTTGGGCACTGCACTGCTTGTACCAAGAGAGACTGCGGGGGAGACGTGTTGTG 1889

QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 1890 CTCGCCACCGCCACCCCTCCGGGCTCCGTCAGTCGTGCCATCCCAATCGAGAGGTT 1949
 QY 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyValAlaIleProLeuGluAlaIle 360
 DB 1950 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAGATATC 2009
 QY 361 LysGlyValArgHisLeuIlePheCysHisSerIleValSerCysAspGluAlaAla 380
 DB 2010 AAGGGGGGGAGACATCTCATCTTCTGTCATTCAGAGAGAGAGTGCAGAGATCGCCGCA 2069
 QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2070 AAGCTGTGCTGATGGGCATCATATGCGGTGGCTACACCGCGTCTTGACGTGCTGCTC 2129
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuThrThrGlyPheThr 420
 DB 2130 ATCCCGACCAAGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2189
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 2190 GGGGACTTGCATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2249
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2250 CTGACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 2309
 QY 461 GlnArgArgGlyValArgThrGlyValArgGlyValArgGlyValArgGlyValArgGly 480
 DB 2310 CAACGTGCGGGGAGACCTGGCAGGGGAGACCTGGCAGGGGAGACCTGGCAGGGGAGAC 2369
 QY 481 GlnArgProSerGlyMetPheAspSerValLeuGlyCysTyrArgAlaGlyCys 500
 DB 2370 GAGGGCCCTCCCGCATGTCGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 2429
 QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 2430 GCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2489
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGlyValPheThrGlyLeu 540
 DB 2490 CCGGGGCTCCCGGTGTCAGACCATCTTGAATTTGGAGGGGCTTTTACAGGGGCTC 2549
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValLeuLeuProTyr 560
 DB 2550 ACTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2609
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerTyrAsp 580
 DB 2610 CTGGTAGCTTACCAAGCCACCGTGTGCTTACGGGCTTACGGGCTTACGGGCTTACGGG 2669
 QY 581 GlnMetTyrPyrCysLeuIleArgLeuValProThrLeuHisGlyProThrProLeuLeu 600
 DB 2670 CAGATGTGAAGT 2729
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrIleValIle 620
 DB 2730 TACAGACTGGGCGGTGTTCAGAAATGAATACACCTTCAGCAGCAGCAGCAGCAGCAG 2789
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrPheValIleValIle 640
 DB 2790 ATGACATGATGTCGCGCGACCTGGAGGTGTCTACGACGACCTGGAGGTGTCTGGAG 2849
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValIleVal 660
 DB 2850 GTCTGTGCTGCTTGTGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2909
 QY 661 IleValLeuSerGlyValProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 2910 GTGTCTTGT 2969

QY 681 AspGluMetGluGluCys 686
 DB 2970 GATGAGATGAGAGATGTC 2987
 RESULT 4
 LOCUS 109328 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 8 from Patent WO 8904669.
 ACCESSION 109328
 VERSION 109328.1 GI:587963
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 5360)
 AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
 JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
 FEATURES
 source Location/Qualifiers
 1..5360
 /organism="Unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,74e-200 Length: 5360
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Gaps: 0
 DB: 6
 US-09-930-591-2 (1-686) x 109328 (1-5360)
 QY 1 MetAlaProIlePheThrAlaTyrAlaGlnIleThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 930 CTGGGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 989
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 990 AGCTTAACCTGGCGGAGCAAAAACCAAGTGGAGGGTCAAGTTGTGTCACTGCT 1049
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValIleHisGlyAla 60
 DB 1050 GCCCAACCTTCTCTGCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 1110 GGAAACAGACCAATCGCGTACCCAGAGGCTGTGTATCCAGATGATACCAATGTAGAC 1169
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1170 CAGAGCTTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 1229
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 1230 GGCTCTCGGACCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 QY 121 GlyAspGlyValArgGlySerLeuLeuSerProAlaArgProIleSerTyrIleValSerSer 140
 DB 1290 GGTGATAGCAGGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1349
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValAlaGlyIlePheArgAlaAlaValCys 160
 DB 1350 GGGGGTCCGCTGT 1409
 QY 161 ThrArgGlyValAlaValAlaValAlaAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 1410 ACCCGTGAAGTGGCTAAGCGGAGCTTATCCCTGTGAAGAACCTAAGACCAACCATG 1469
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 1470 AGGTCCCGGTGTTCAGGATTAATCTCTTCCACCAAGTAGTGCCCGCAGAGTTCCAGGTG 1529

934
 1213
 934
 273
 1034
 934-1213>100

QY 201 AAlhHsleuHsAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 1530 GCTCACTCCATGCTCCACAGGAGCGCAAAAAGCCAAAGGTCCCGCTGATATGCA 1589
 QY 221 AAlGlnGlyTyrLysValLeuValLeuAspProSerValAlaAlaThrMetGlyPheGly 240
 DB 1590 GCTCAGGGCTATAGAGGTGCTAGTACTACCCCTCTGTTCTGTCACACCTGGGCTTTGGT 1649
 QY 241 AlaTyrMetSerLysAlaHsGlyLysLeuAspProAsnLysLeuThrGlyValArgThrIle 260
 DB 1650 GCTTACATGTCACAGGCTCATGGATGATCTTAACATCAAGACCGGGGAGAAACAAT 1709
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 1710 ACCACTGGCAGCCCACTACGACTCTCACTACGCAAGTCTCTGCGACGGGGGTGC 1769
 QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHsSerThrAspAlaThrSer 300
 DB 1770 TCGGGGGGGCTTATGACATTAATTTGTCACAGAGTCCACTCCAGATGCCACATCC 1829
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValArgLeuThrVal 320
 DB 1830 ATCTTGGGCATCGGACATGCTCTTACCAAGCAGAGACTGGGGGCGAGACTGGTTGTG 1889
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHsProAsnIleGluGluVal 340
 DB 1890 CTCGCCACCGCCACCTCCGGGCTCCGTCACATGCCCCATCCCAACATGAGAGGTT 1949
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluValIle 360
 DB 1950 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGTATCCCTCCGAAAGTATC 2009
 QY 361 LysGlyGlyValArgHsLeuIlePheCysHsSerLysLysCysAspGluLeuAlaAla 380
 DB 2010 AAGGGGGGAGACATCTCATCTCTCTGTCATCAAGAAGAGTCCGAGAACTCGCGCA 2069
 QY 381 LysLeuValAlaLeuGlyValAspAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2070 AAGCTGTGCGCATTTGGGCAATGATGCCGTGCTACACCGCGCTTGAAGTCTCCGT 2129
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 2130 ATCCGACACGAGGGGATGTTGCTGCTGCAACCGATGCCCTCATGACCGGCTATACC 2189
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrValAspPheSer 440
 DB 2190 GCGGACTTCGACTCGGTATGACATGCAATACGTGTGCACCCAGACAGTGCATTTACG 2249
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2250 CTGACCTTACCTTACCATTTGAGCAATCAAGCTCCCGAGATGCTGTCTCCGCACT 2309
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 2310 CAACGTCGGGGGAGAGACTGCGAGGGGAAAGCAGGCACTACAGATTGTGCGACCGGG 2369
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyTyrAspAlaGlyCys 500
 DB 2370 GAGGCCCTCCCGGATGTCAGTCTGCTCTCTGTGATGCTATGACGACGAGCTGT 2429
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 2430 GCTTGGTATGAGCTACGCGCGCGAGACTACAGTTAGCTACGAGCTATGAAACACC 2489
 QY 521 ProGlyLeuProValCysGlnAspHsLeuGluPheTyrGluGlyValPheThrGlyLeu 540
 DB 2490 CCGGGGCTTCCCGTGTGCGAGACCATCTTGAATTTGGAGAGGGGTCTTTACAGGCTTC 2549
 QY 541 ThrHsIleAspAlaHsPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 2550 ACTCATATATATGAGCCACTTTCTATCCGACAAAGCAGAGTGGGGAAGAACTTCTTAC 2609
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerSerTyrAsp 580

DB 2610 CTGATAGCGTACCAAGCCACGCTGTGGCTAGGGCTCAAGCCCTCCCATCGGGGAC 2669
 QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHsGlyProThrProLeuLeu 600
 DB 2670 CAGATGTGAAGTGTATTTATTCCTCAAGCCCACTTCATGGGCGCAACCCCTGCTA 2729
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHsProValThrLysTyrIle 620
 DB 2730 TACAGCTGGGCGCTGTTCAAGATAAATCACTCGACGCCAGCTACCAAAATACATC 2789
 QY 621 MetThrCysMetSerAlaAspLeuGluValIleThrSerThrTyrPValLeuValGlyGly 640
 DB 2790 ATGACATGATGTCCGGCGCACTGAGAGGTGTCACAGACACCTGGGTCTGTCGGCG 2849
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 2850 GTCTGGCTGCTTTGGCGCGGTATTTGCTGTAAACAGGCTGCGTGTCTATAGTGGGCG 2909
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 2910 GTGCTTTGTCCGGAAAGCCGCAATCACTGACAGGAAGTCTCTACCGAAGTTC 2969
 QY 681 AspGluMetGluGluCys 686
 DB 2970 GATGAGATGGAAGAGTGC 2987
 RESULT 5
 LOCUS 106440 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 54 from Patent EP 0318216.
 ACCESSION 106440
 VERSION 106440.1 GI:590312
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6785)
 AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
 TITLE NAbv diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;
 FEATURES
 source 1..6785
 /organism="Unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,32e-200 Length: 6785
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x 106440 (1-6785)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1203 CTGGCGGCCATCAAGGGTATGCGCCAGACAGAAAGGGGCTCTTACGGTGCATTAATCAC 1262
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 1263 AGCTTAACCTGGCGGCGCAAAAACCAAGTGGAGGGTGAAGTCCAGATGTGTCAACTGCT 1322
 QY 41 AlAGlnThrPheLeuAlaThrCysIleAsnGlyValCysThrThrValTyrHsGlyAla 60
 DB 1323 GCCCAAACTTCTCCGCAACGTGATCAATATGGGTGTGCTGCACTGTCAACAGGGGCC 1382
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 1383 GGAACGAGACATCGGTGTCACCAAGGGTCTGTCTCATCAGATGTATACAAATGTATGAC 1442

QY 81 GlnAspLeuValAlaGlyTrpProAlaProGlnGlyValAaArgSerLeuThrProCysThrCys 100
 Db 1443 CAAGACCTTGCTGGGCTGGCCGCTCCGCAAGGTAGCTCATATGACACCTGCACTTGC 1502
 QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 1503 GGGCTCCCGAACCCTTACCTGGTACAGAGCACCCATGTCATTCCTGGCCGCGCGG 1562
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerThrLeuGlySerSer 140
 Db 1563 GGTATAGCAGGGGACGCTGCTCCGCGCGCCATTTCTTCTTGAAGGCTTCTCG 1622
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1623 GGGGGTCCGCTGTTGTGCCCCGCGGGGACGCGTGGCATATTAGGCGCGGCTGTC 1682
 QY 161 ThrArgGlyValAlaValAspPheIleProValGlySerLeuGluThrThreMet 180
 Db 1683 ACCCGTGAAGTGGCTAAGCGGTGGACTTATCCCTGTGAGAACCTTAGAGAACCATG 1742
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerThrGlnVal 200
 Db 1743 AGGTCCCGGCTTTCACGATACCTCTCTCCACAGTAGTCCCGACAGCTTCCAGGTG 1802
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaValAla 220
 Db 1803 GCTACCTTCATGCTCCACAGGACGGGGAAGAACCAAGTCCCGGTGATGATGCA 1862
 QY 221 AlaGlnGlyThrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1863 GCTCAGGCTATAGAGGTAGTACTCAACCCCTGTGTGTGTCAGAACCTGGCTTGTGT 1922
 QY 241 AlaThrMetSerLysValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 1923 GCTTACATGTCACAGGCTCATGGATGATCTTACATCAGACCGGGGTGAGAACATTT 1982
 QY 261 ThrThrGlySerProIleThrThrSerThrThrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1983 ACCACCTGGACGCCCATCAGTACTCCACCTAGCGGAAGTTCCTTGGCGAGGGGGTGC 2042
 QY 281 SerGlyGlyAlaThrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
 Db 2043 TCGGGGGGCGCTTATGACATATATTTGTGACAGAGGCCACTCCAGGATGCCACATCC 2102
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2103 ATCTTGGGACATCGGACATGCTCTTGACCAAGACAGACTGGGGGGGAGAGCTGTGTG 2162
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2163 CTGGCCACCGGACCCCTCGGGGCTCCGTACTGTGCCCCATCCCAACATCGAGGAGTT 2222
 QY 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2223 GCTCTGTCACCAACCGAGAGATCTCTTTTACCGGCAAGGCTATCCCCCTCGAAGTATC 2282
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerThrLysValCysAspGlyLeuAlaAla 380
 Db 2283 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAGAGAGAGTCCGAGAACTGCGCGCA 2342
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaValThrArgGlyLeuAspValSerVal 400
 Db 2343 AAGGTGGTGGCATTTGGGACATCAAGCCGTGGCTCACTACCGCGGTCTTGAAGTCTCGTC 2402
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2403 ATCCCGACAGCGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2462
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGluThrValAspPheSer 440
 Db 2463 GGGACATTTGACCTCGGATAGACTGCATATGCTGTGTGTGTGTGTGTGTGTGTGTGT 2522
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 2523 CTTCACCTTACCTTTCACCATTTAGACATCACCTCCCGAGATGCTGTCTCCGCACT 2582
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleLeuThrArgPheValAlaProGly 480
 Db 2583 CAAGCTGGGGGAGAGCTGGACAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2642
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysThrAspAlaGlyCys 500
 Db 2643 GAGCGCCCTCCGCGCATGTTTCAGCTGCTGCTCTGTAGTATGATGACGAGGCTGT 2702
 QY 501 AlaThrThrGlyLeuThrProAlaGluThrThrValArgLeuArgAlaIleMetAsnThr 520
 Db 2703 GCTTGGATAGCTCAACGCGCGGACAGACTACAGTTAGGCTACGAGGCTACGACACCC 2762
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 2763 CCGGGGCTTCCCTGTGGCCAGAGACATCTTGAATTTGGAGGGCGCTTTTACAGGCTTC 2822
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProThr 560
 Db 2823 ACTCATATAGATGCCCATTTCTATCCAGACAAAGAGGTGGGAGAACCTTCTTAC 2882
 QY 561 LeuValAlaThrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 Db 2883 CTGGTAGCTACCAAGACCGGTGGCTAGGGCTCAAGCCCTCCCATCGTGGGAC 2942
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 2943 CAGATGTGAAGTGTGATTTGCTCCCTCAAGCCCACTTCATGAGGCCCAACCCCTGTGCTA 3002
 QY 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisAspProValThrLysThr 620
 Db 3003 TACAGACTGGCGCCTGTTCAGATGAATACCTTACAGACCAAGTCCCAATATCATC 3062
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 3063 ATGACATGACATGTGCGCGGACCTGAGGTGTGACAGACGACCTGGGGTGTGCTGTGGCGGC 3122
 QY 641 ValLeuAlaAlaLeuAlaAlaValThrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 3123 GTCCGTGCTGCTTGTGGCGGATGCTGTCTGTCAACAGGCTGCGTGTGATAGTGGGAGG 3182
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuThrArgGluPhe 680
 Db 3183 GTGCTGTGTCGGGAGAGCGGACATCATCTGACAGGAGATCTCTTACAGAGTTC 3242
 QY 681 AspGluMetGluGluCys 686
 Db 3243 GATGAGATGGAAGAGTGC 3260

RESULT 6
 LOCUS 109329 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 10 from Patent WO 8904669.
 ACCESSION 109329
 VERSION 109329.1 GI:587964
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES 1 (bases 1 to 6785)
 Houghton, M., Choo, O.-K., and Kuo, G.
 Patent: WO 8904669-A 10 01-JUN-1989.
 JOURNAL Location/Qualifiers
 FEATURES
 source /mol_type="unknown"
 ORIGIN
 Alignment Scores: 2,32e-200 Length: 6785
 Pred. No.: 3574.00 Matches: 672
 Score:


```

RESULT 7
ARI18696 7310 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 74 from patent US 6150087.
DEFINITION ARI18696
ACCESSION ARI18696
VERSION ARI18696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 7310)
TITLE Chien, D. Y.
JOURNAL NMBV diagnostics and vaccines
FEATURES Patent: US 6150087-A 74 21-NOV-2000;
          Location/Qualifiers
          source 1..7310
                /organism="unknown"
                /mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 2.54e-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x ARI18696 (1-7310)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATTCACGGGCTACGCCACGACGACGAGGGGCTCTCTAGGCTGATATACCC 1787

QY 21 SerLeuThrGlyArgAspGlyAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCTTAACGTGGCCGCGACAAACCAAGTGGAGGGTGAAGTCCAGATTGTCAACTGCT 1847

QY 41 AlAGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1848 GCCCAAACTTCTCGGACAGTGCATCAATGGGGTGTGCTGACTGTCTACCAACGGGGCC 1907

QY 61 GlyThrArgThrIleAlaSerProGlyGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGACATCGGCTCACCCAGAGGCTCTCTCATCCAGATGTATACCAATGTAGAC 1967

QY 81 GlnAspLeuValGlyTTPProAlaProGlnGlyValAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGAGGCTGGCCGCTCCGCAAGGTAGCGCTCATTGACACCCCTGCATTGC 2027

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 2028 GGCTCTCTGAGACTTTCATCTGTGCACGAGGACCGCATGTCTCCGTGCGCGCGCGG 2087

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlyGlySerSer 140
Db 2088 GGTATATGACAGGGGACGCTGCTCTCGCCCGGCCCATTTCTTCTTAAGAAAGCTCTTCG 2147

QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTTGTGCCCGCGGGGACGCGTGGGCAATATTAGGGCGCGGCTGTGC 2207

QY 161 ThrArgGlyValAlaAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCGTGTAGGTGGTAAAGCGGTGACTTTATCTCTGTGAGAACTTAGAGACAACATG 2267

QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTCAAGGATACCTCTCTCCACCAAGTAGTGCCCAAGGCTTCAAGTGTG 2327

QY 201 AlaHisLeuHisAlaProThrGlySerGlyCysSerThrIysValProAlaAlaTyrAla 220

```

```

Db 2328 GCTCACTTCAGTCTCCACAGGACGGCGCAAAAGCAACCAAGTCCCGCTGCATATGCA 2387
QY 221 AlAGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATTAAGGTGCTAGTACTCAACCCCTCTGTGCTGCACACACTGGGCTTTGCT 2447

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGACCGGGGAGAAACAAT 2507

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCATTCACGATCTCACCTTACGGCAGATCTCTTGGCCGAGCGGGGTGC 2567

QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGGCGCTTATGATCATATATTTTGGAGAGTGCACATCCACGATGCCACATCC 2627

QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTGGGCATCGGACGTGCTTGACCAAGCAGAGACTCGGGGGGAGACTGTGTGTG 2687

QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCCACCCCTCGGGCTCGCTCACTGTGCCCCATCCCAACATCGAGAGAGTT 2747

QY 341 AlaLeuSerThrThrArgGlyIleIleProPheTyrGlyIysAlaIleProLeuGlnAlaIle 360
Db 2748 GCTCTGTCCACCAACCGAGAGATCTCTTTTACGGCAAGGCTATCCCTTCGAAATATC 2807

QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCATCTTCTGTGCATTAAGAAGAGAGCGACGAACTCGCGCA 2867

QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGTCTCATTTGGACATCAATCGGTGGCTTACTACCGCGCTTGAACGTGTCCGCT 2927

QY 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCGACACGCGGCGATGTGTGCTGTGGCAACGATGCCCTTCATGACCGGCTATACC 2987

QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GGGACCTTGACCTCGGATGATGACTGCATACGTGTGTGCCACCAAGTCGATTTCCAGC 3047

QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCTTACCTTTCACCATTTGAGACATACACCTCCCGAGATGCTGTCTCCGCACT 3107

QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACGTGGGGGAGAGTGGACAGGGGAGGAAACGAGCATTTACGATTTGTGGACCGGGG 3167

QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATTTGCACTGCTCGTCTGTGATGCTATGACGAGGCTGT 3227

QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGTGATGATGCTACGCGCGGAGACTACAGTTAAGGTACGAGGTTCATGACACCC 3287

QY 521 ProGlyLeuProValCysGlnAspHisIleGluIlePheThrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCGCTGTGCCAGAGACATCTTGAAATTTTGGAGGGCGCTTTTACAGGCTTC 3347

QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATATATCCCATCTTCTATCCCAAGACAAAGAGAGTGGGAGAACTTTCCTTAC 3407

QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrPasp 580
Db 3408 CTGTGTAGCTATCAACGACCGGTGTGGCTAGGGCTCAAGCCCTCCCTCATGTGTGGAC 3467

```

QY 581 GImetrrpLysCySleuileuLeuysProthLeuHISGLYProthProleuLeu 600
DB 3468 CAGATGGGAAGGTTGATTGCTCAAGCCCACTCCATGGGCCAACACCCCTTGCTA 3527
QY 601 TTAATGleuGlyAlaValGlnAengluValThrLeuThriSProValThriLysTyrIle 620
DB 3528 TACAGACTGGGCGCTGTTCAGATGAATATCAACCTTCAGCGCACTCCATCAAAATCAATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 3588 ATGACATGATGTCGGCGGACCTGGAGGTGTCAAGACACCTGGGTGCTCGTTGGCGGG 3647
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 3648 GTCCGTGGCTGCTTTGGCCGCGCTATTGCTGTCAACAGCGCTGCGGTCAATGAGGCAAG 3707
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 3708 GTCCGTCTGTCCGGGAAGCCGGCAATCATCTGACAGGAGGAGTCTCTACCGAGAGTTC 3767
QY 681 AspGluMetGluGluCys 686
DB 3768 GATGAGATGAAAGAGTGC 3785

RESULT 8
109331 7310 bp DNA linear PAT 02-DEC-1994
LOCUS Sequence 15 from Patent WO 8904669.
DEFINITION 109331
ACCESSION 109331 GI:587966
VERSION 109331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7310)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES
source 1..7310
/organism="unknown"
/mol_type="unassigned DNA"

ALIGNMENT Scores: 2,54e-200 Length: 7310
Pred. No.: 3574.00 Matches: 672
Score: 99.42% Conservative: 10
Percent Similarity: 97.96% Mismatches: 4
Best Local Similarity: 98.78% Indels: 0
Query Match: 6 Gaps: 0
DB: US-09-930-591-2 (1-686) x 109331 (1-7310)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1728 CTGGCGGCATCAACGGGTACGCGCCAGACAAAGGCGCTCTCTAGGGTGCATATCAACC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 1788 AGCTTAATCTGGCGGACAAACAAAGTGAAGGTGAGGTCCAGATTGTCTCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 1848 GCCCAACCTTCCCTGGGACAGTGCATCATGGGTGTGGAGCTGTCTACCGAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1908 GGAACGAGACATCGGTCTCAACCAAGGTCTGTCTATCAAGATGATACCATGTAGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaIleArgSerLeuThrProCysTrpCys 100
DB 1968 CAAGACCTGTGGGTGGCCGCTCCGACAGTATGCGGTCTATGACACCTGCACCTTGC 2027

QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 2028 GGCTCCCTCGAGACCTTATCTGTGTGTCAGAGCAAGCCGATCTATCTCCCGTGGCCGGCG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 2088 GGTGATAGCAGGGGCACTGCTGTGCGCCCGGCCATTTCTTCTTGAAGGCTCTCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 2148 GGGGTCCGCTGTGTGTGCGCCGGGACAGCGGTGACATATTAGGCGCGGGTGTGC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 2208 ACCGTGAGATGGCTTAAGCGCGGTGAGCTTATCTCCCTGGAGAACTTGAAGACAAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 2268 AGGTCCCGGTGTTCACGATTAATCTCTCCACACAGTATGCCCCAGAGCTTCAGAGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 2328 GCTCACCTTCATGCTCCACAGCAGCGGCAAAAGCAACAAGTCCGGCTGCATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 2388 GCTCAGGGCTATTAAGTGTGATGACTCAACCCCTGTGTGCAACCTGGGCTTGTGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 2448 GCTTACATGTCCAAAGGCTCATGAGATGCATCTTAACATCAAGACCGGGGTGAAACAATT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 2508 ACCACGGAGAGCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2567
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 2568 TCGGGGGGGCTTATACATTAATTTGTGACAGTGCATCCACGAGATGCCACATCC 2627
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 2628 ATCTTGGGATGCGGACCTGCTTACCAAGACAGACGCGGGCGGAGCTGCTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 2688 CTCGCCACCGCCACCCCTCGGGCTCGTCACTGTCCCACTCCCAACATCCAGAGAGTT 2747
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGCGCAAGGCTATCCCTCGAATGATTC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 2808 AAGGGGGGAGACATCTTCATCTCTGATCAATCAAGAAAGAGGACCAATCCGCCGCA 2867
QY 381 LysLeuValAlaLeuGlyAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 2868 AAGCTGGTGGCATTTGGGATCAATGCGGTGCTTACACCGCGGTCTTGAAGTGTGCTGCT 2927
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 2928 ATCCGACCAACGGCGAGT 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 2988 GGCGACTTGCAGCTGGGTGATGACTGCATACGATGTGTCAACCAAGTGCATTTTCAGC 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 3048 CTTGACCTTACCTTACCATTCATGAGACATCACTCCCGCGAGATGTGTGTGTGTGTGTGT 3107

Qy	461	GlhAtrgRgGlyAtrgThrglyAtrgGlyAtrgProglyAtrgThrgpneValAtrgProgly	480
Db	3108	CAACGTCGGGGCAGAGCTGCGAGGGGAGACGAGCATCTACAGATTGTGGACCGGGG	3167
Qy	481	GlhAtrgProserGlymetPheapserSerValleucysGlyCysThrgpneAtrgAtrgCys	500
Db	3168	GAGCGCCCCCTCCGAGCATGTCGACTGCTCCGCTCTGTAGTGTATGACGAGCGCTGT	3227
Qy	501	AlatrPrpGlyleuThrProalaglThrThrValAtrgLeuAtrgAtrgAtrgMetAtrgThr	520
Db	3228	GCTTGGATAGCTCAGCGCCCGGAGACATACGTTAGCTACGAGGCTACATGAAACACC	3287
Qy	521	ProglyleuProValCysGlnaAtrgAtrgLeuThrgpneThrgpneThrgpneThrgpne	540
Db	3288	CCGGGGGCTCCCGTGTCCGAGACCATCTGAAATTTTGGAGGGCGCTTTACAGGGCTTC	3347
Qy	541	ThhAtrgLeuAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	560
Db	3348	ACTCATATAGATGCCCATCTTCTATCCCAACAAAGAGTGGGAGAACTTCTCTAC	3407
Qy	561	LeuValAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	580
Db	3408	CTGTAGACCTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCTCCCATCTGTGGAC	3467
Qy	581	GlmMetTrpLysCysleuAtrgLeuAtrgLeuAtrgLeuAtrgLeuAtrgLeuAtrgLeu	600
Db	3468	CAGATGTGAAGTGTGATTGCTGCTCAAGCCCACTCATGGGCAACACCCCTCTCTA	3527
Qy	601	TyrAtrgLeuGlyAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	620
Db	3528	TACAGATGGGCGCTGTTCAGAAATGAAATACCTTGACGACCAACCAATATATATC	3587
Qy	621	MetThrCysMetSerAlaAtrgLeuAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	640
Db	3588	ATGATCATGATGTCGCGGACCTGTGAGTGTCTACGAGCACTGGGCTCTGTGGCGGC	3647
Qy	641	ValleuAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	660
Db	3648	GTCCTGGCTGCTTGGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3707
Qy	661	IleValleuSerGlyAtrgProAlaIleIleProAtrgAtrgAtrgAtrgAtrgAtrgAtrg	680
Db	3708	GTCGTCTTGTCCGGGAAGCCGCAATCATCTGACAGGAAAGTCTCTACGAGGATTC	3767
Qy	681	AspGluMetGluGluCys	686
Db	3768	GATGATGATGAAAGATGC	3785

ORIGIN

RESULT 9
HPCPOLYP 7310 bp ss-RNA linear VRL 02-AUG-1993
LOCUS Hepatitis C virus polyprotein gene, partial cds.
DEFINITION
ACCESSION M32084
VERSION M32084.1 GI:329875
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
viral; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
1 (bases 1 to 7310)
Choo, Q.-L., Richman, K. and Han, J.
The nucleotide sequence of the Hepatitis C viral genome
Unpublished (1990)
Original source text: Hepatitis C virus, CDNA to viral RNA, clones
K9-1 through 15e, isolated from chimpanzee (individual 910) blood
plasma.
Draft entry and printed sequence for [1] kindly submitted by
M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street,
Emeryville CA 94608.

CDS

Qy	1	MetAlaProIleThrAlaTrpAlaGlnThrAtrgGlyLeuLeuGlyCysIleIleThr	20
Db	1728	CTGGCCCATTCACGCGCTACGCGCCACGACGAGGGGCTCTCTAGGCTGCAATATCACC	1787
Qy	21	SerLeuThrGlyAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	40
Db	1788	AGCCTAATCTGGCGGAGCAAAACCAAGTGAAGGTGAGGTCTCAGATTGTGTCAACTGCT	1847
Qy	41	AlaGlnThrPheLeuAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrgAtrg	60
Db	1848	GCCCAAACTTCTGCGACATGATCAATGAGGGGTGTGCTGAGCTGTCTACGAGGGGCC	1907

Alignment Scores:

Pred.:	2.54e-200	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	14	Gaps:	0

US-09-930-591-2 (1-686) x HPCPOLYP (1-7310)

/db_xref="taxon:11103"
<1..>7310
/note="polyprotein"
/codon_start=3
/protein_id="AA045677.1"
/db_xref="GI:329875"

/translation="GCERLASCRLPTDQCKGPISTANGSGPPDRPYCMHPPKRC
GIVPAKSCGVPVCFPTSPVAVGTTDRSAPATYSWENLDTVFNNTPLCNMFGC
TMMNSYGFPRKVCAPPCVIGAGNNMTLCPCTCFRKHGPRATYSCSGEMPTFRLAVD
YPRLWHNPCTINVTYFKIRMYGVJEHLCAAMTNRGRCDELDSELSPLALD
TTOMOVLPSCFTTLPALSTGHLHONIVDOVLYGVGSSISMAIKWEVYVLPFLT
ADARVSCILMMHLLISQARALENYIILNAAISLACHGIVSLPVEPCSMVYKGVV
GAVTYTGMPPLMLLALPQRAYALDITVAASCGGVVLGMLATLSPPYKYSIC
LWMLQYELFRVEAQLHWIPLNVRGRAPVILMCAVHTVFDITKLIAVFGPLM
IIQASLLKTPYEVYVGLRFLCAAKMKGHVQVVIKLGALITCTYVNHLPRLD
MANGRDLAVANPEVVFQSMETKLTWADPACDIIINGLFSVARGELILGAD
GVNSKGMRLAPITAAQDTRGLGCIILSLRGDNGVGEQVISTAAQPLATCI
NGCVTVYHAGRTTASPKPVIQMTYVDDDLVGMPPQGSRLTPTCCSSDLTL
VTRHADVIVRRRGDSLSLSPRISTYKSGGGLPDPAGHAGVIFRAAVTRGVA
KAVDFIPENLETTMRSVFTDNSPVPVQSFQVNHLPAGTSGKSTKVPAAVAAOG
YKVLVNPVAAITLFGAAMSRAHGIDPMIRGVRTTGTSPITVETGFLADGGS
GGAVIDIICDECHSDATSIIGITVLDQAEAGALVLAATPSPGVVPHNIE
VALSTTGEIIPYKGAILEVIGKGRHLICHSKKKDELAALKVAGINAVAYRGD
VSVITPSGVVVAIDALMTGTGDDSDYIDCTCYTQVYDSDPTFTTETITLSD
AVRTQRRRTGKPGIYRFAVGERPFGMSVLCEDYAGCAMTELTPELTIVR
LAVYMTPGILPVODHLEFWEVGTGLTHIDAFLSQTKOSGNLPLVAVOATVCR
AQPSPSPDQMKCLIRLKLPTLHGPPLIRLGAVNEITLTPVTKYIMTCSADLE
VYTSVTLVGVVLAALAAVCLSTGCVVIVGRVVLGSKPAIIPRELVREFDMEBS
CHLPTYEGQMMALAEQFKALGLITQASQAVIARAVOTMMOKLETFFAKHMMFTS
GIQYLGSLTLPENPAISLMAFTAVTPTLTSOTLLENILOGVAAQLAAGAAITA
PVGAGLAGAIGVSGKAVLIDLAGYAGVAGALVAFKIMSGEVSTEDLVNLEPAI
LSFGALVGVVCAAILRHVYGRBEGAVOMNRLIAPASGNHVSPTHYVESDAARV
TALLSLVYQTLRLRHOMISECTPCSGSWRLDMDIICEVLSPFKMLKAKMPO
LGEIPVSCORGKGYWRVDMHTRCHGASLETGKNGKTRIVPRTCRNWMSGF
PINAATYTGCTPLPAPNTPALMRVASEEYVIRVQGDHYVYGMPTDMLKCCOVS
PEFTTLDGRLHRAFPCKPLAREVSRVGHIERPVSCDLPCEBEPYAVUTLSMT
DPSHTIAEAGRLHRAFPCKPLAREVSRVGHIERPVSCDLPCEBEPYAVUTLSMT
OWENGNITRVESNEKVVILDSPDLAEEDERIVPAELIRKSRFAQLPMWARD
YNPPLVETWKEDEYEPVVGCPILPPKSPVPPPKKTTVLTSETLSTALMLTR
SFSSSTSGITGNTTSSBPASGCPSPDSASVSMPLDEGPDPDLSDDSMSPV
SSRANMDEVCSSMSYSGALVTPCAAREOKLPIALNSILRHNNIVYSTSRAC
OROKKTFRLQVDSHYDVLKEVNAASRYKANLISFEACSLTPPHSAKSKFGG
ACQVRHAKRAVTHINSWKDLLEDVNTIIVTTIMKNVFCVQAKGKKKAPRLVIF
PDLGVAVCEKMAIYDVYTKLPLAVMSVSGVFOARVEFLQAWKSKTTPGFSFD
TRCFDSTVIESDIRTEBALYOCDDLPQARVAKSLTERLYVGGPTNSRGCGAR
CRASGLVLTSSGNTLTCYIKARACGAAGLQDCTMVCGBDLVVICESGVEDASL
RAETEMTRYSAPEPDPPEYDLEITGSCSSVSAHNGAGRYVYTLRPTPLAR
AAWETARHTPVNSWLGNTIMFAPTLMARMLMTHFVSVLARQDLQALDCEIYGYC
SIEPDLPIPIORL"

QY	61	GLYThrArgThrIleAlaSerProLysGlyProValIleGlnMetCysThrAspValAsp	80
Db	1908	GGAAcGAGAGCAATCGCCTCAcCAAGAGGTCCTGTCAcGAAAGTAAcCAAGTATGAC	1967
QY	81	GLnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	1968	CAAGACCTTGCGGCTGGCCCGCTCCGAAAGTAGCGCCTATTGACACCCCTGCACCTTGC	2027
QY	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
Db	2028	GGCTTCCTCGACCTTTACTGTGTCAcGAGGACACCGCGATGTCACTTCCCGTGGCGGGG	2087
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	2088	GGTAGACAGAGGGGcAGcCTGTGTGTGCCCGGCCCATTTCTCTACTTGAAGGCTCTCG	2147
QY	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
Db	2148	GGGGGTCCGCTGTGTGTGCCCCGGGGGACCGCCGAGCATATTTAGGGCCGCGGTGTGC	2207
QY	161	ThrArgGlyValAlaValAlaValAspPheIleProValGlySerLeuGluThrThreMet	180
Db	2208	ACCGCTGAGGGCTCAAGCGGTGACTTTATCCCTGTGGAGAACTTAGAGACATAG	2267
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrAlaVal	200
Db	2268	AGGTCCCGGCTTCAcCGATACCTCTCTCCACAGATAGTGCcCAAGGCTTCAGAGTG	2327
QY	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaArgTyrAla	220
Db	2328	GCTCAcCTTCACATGCTCCACAGGcAGGcGcCAAAACACCAAGTCCGAGCTGCATATCA	2387
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	2368	GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTGTGTGTGCTGCAACCTGGGCTTTGGT	2447
QY	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	2448	GCTTCAcATGTCcCAAGGCTCATGAGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT	2507
QY	261	ThrThrArgLysSerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	2508	ACCATCTGGAGGCCCAcATCAcGTAcTTCACATTCAGGAAAGTCTCTTCCGAGCGGGGTGC	2567
QY	281	SerGlyGlyAlaTyrAspIleIleIleCysAspArgLysHisSerThrAspAlaThrSer	300
Db	2568	TGGGGGGGGCTTATGACATATATATTTGTGACAGTGCcCATCTCCAGCGATCCCAATCC	2627
QY	301	IleLeuGlyTyrLeuGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	2628	ATCTTGGGcATGGcCATGCTCTTGTGACCAAGAGAGAcTGGCGGGGcAGAcTGTGTGTG	2687
QY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	2688	CTCGcCAcCGcCAcCCCTCGGGCTCGTCAcTGTGGCCCATCCCAcATCGAGGAGGTT	2747
QY	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	2748	GCTCTGTCCACCAcCGAGAGATCCCTTTTTCAGGAGAGCATCCCTCTGAAcGTATTC	2807
QY	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	2808	AAAGGGGGGAGACATCTCATCTTCTGTCACTTAAGAGAGAAAGTGCAGCAATCGCGCGA	2867
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	2868	AAAGCTGTGcATGTGGcCATATGCcCGTGGcCTACTACCGGAGTCTTCAcGTGTCCGTC	2927
QY	401	IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr	420
Db	2928	ATTCcCAcCAcGcGGcCATGTGTGTGCTGTGGGcCAcCGAAGcCTCATGAcCGGCTATACC	2987
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440

Accession	Source	Sequence	Length	Start	End	Strand	Feature
D8288	GenBank	GGCGACTTCGACTCGGATGATGACTGCATACCGTATACCGGACAGTCGATTTCAGC	3047	1	3047	+	
D441	GenBank	LeuAspProThrPheThrIleGluThrIleThiLeuProGlnAspAlaValSerArgThr	460	441	900	+	
D3048	GenBank	CTTACACCTTCACCTTCACCATGAGACATACACGCTCCCCAGAGTGTGTCTCCGGACT	3107	3048	6155	+	
D461	GenBank	GlnArgArgGlyArgThrGlyArgGlyValSerProGlyIleTyrArgPheValAlaProGly	480	461	940	+	
D3108	GenBank	CACGCTGGGGCGAGGACTGGCCAGGGGGAGCCAGGACCTTACAGATTGTGGACCGGGG	3167	3108	6275	+	
D481	GenBank	GlnArgProSerGlyMetPheAspSerValLeuGlyGluGlyTyrAspAlaGlyCys	500	481	980	+	
D3168	GenBank	GAGGGCCCTCCGGCAGATGGTTCGACTCGTCCGCTCTGTAGTGCATGACGAGCGTGT	3227	3168	6495	+	
D501	GenBank	AlaThrPyrGluMetThrProAlaGlnThrThrValArgLeuArgAlaTyrMetAsnThr	520	501	1020	+	
D3228	GenBank	GCTTGGATGAGCTTCACGCCCGCCGAGACTACAGTTAGCTACGAGCGTACATGAAACCC	3287	3228	6515	+	
D521	GenBank	ProGlyLeuProValCysGlnAspHisLeuGlnPheThrGluGlyValPheThrGlyLeu	540	521	1060	+	
D3288	GenBank	CCGGGGGCTCCCGTGTGCAGAGACCATCTTGAATTTGGGAGGGCGCTTTACAGGGCTC	3347	3288	6665	+	
D541	GenBank	ThrIleAlaAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValMetLeuProTyr	560	541	1080	+	
D3348	GenBank	ACTCATATGATGAGCCCACTTTCTATCCGACCAAGACGAGATGGGAGAACCTTCCCTAC	3407	3348	6815	+	
D561	GenBank	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp	580	561	1100	+	
D3408	GenBank	CTGTATGGGTACCAAGCACCGCTGTGGGCTACAGGCTCCACCCCTCCCTCGTGGGAC	3467	3408	6975	+	
D581	GenBank	GlnMetTyrPlyCysLeuIleArgLeuLeuProThrLeuHisGlyProThrProLeuLeu	600	581	1120	+	
D3468	GenBank	CAGATGTGGAGTGTGTGATTGGCTCAAGGCCCACTCCATGGGCGCAACCCCTGTCTA	3527	3468	7135	+	
D601	GenBank	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleTyrIle	620	601	1140	+	
D3528	GenBank	TACAGACTGGGCGCTGTTCACAGATGAATACCCCTGACGACCCAGTACCAATATACATC	3587	3528	7295	+	
D621	GenBank	MetThrCysMetSerAlaAspLeuGluValAlaThrSerThrTyrPvalLeuValGlyGly	640	621	1160	+	
D3588	GenBank	ATGACATCATGTCGCGCCGACCTGAGGTCCTCCACGACACCTGGGAGCTCGTGGCGGC	3647	3588	7455	+	
D641	GenBank	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValAlaIleValGlyArg	660	641	1180	+	
D3648	GenBank	GTCCTGGGTGCTTGGCCGCGTATTGCTGTCCACAGGCTCGTGTGTCATAGTGGCAGG	3707	3648	7615	+	
D661	GenBank	IleValLeuSerGlyIlyProAlaIleIleLeuProAspArgGluValLeuTyrArgGluPhe	680	661	1200	+	
D3708	GenBank	GTCGCTTGTCCGGGAACCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC	3767	3708	7785	+	
D681	GenBank	AspGluMetGluGluCys	686	681	1206	+	
D3768	GenBank	GATGAGATGAGAAAGTGC	3785	3768	7873	+	
LOCUS	108294	108294	9185 bp	DNA	linear	PAT 02-DEC-1994	
DEFINITION	Sequence 1 from Patent EP 0388232.						
ACCESSION	108294						
VERSION	108294.1	GI:588994					
KEYWORDS							
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 9185)						
AUTHORS	Houghton,M., Choo,Q.-L. and Kuo,G.						
TITLE	NANBV diagnostics and vaccines						
JOURNAL	Patent: EP 0388232-A1 19-SEP-1990;						
FEATURES	Location/Qualifiers						
	1..9185						
	/organism="unknown"						
	/mol_type="unassigned DNA"						

ORIGIN

Alignment Scores:

Pred. No.: 3,366-200
Score: 3574.00
Percent Similarity: 99.42%
Best Local Similarity: 97.96%
Query Match: 98.78%

Length: 9185
Matches: 672
Conservative: 10
Mismatch: 4
Indels: 0
Gaps: 0

US-09-930-591-2 (1-686) x 108294 (1-9185)

QY 1 MetAlaProIleThraIaIaYrAlaGlnGlnThraGlyLeuLeuGlyCysIleIleThr 20
DB CTGGCGCCCATTCACGGCGTACGCCACAGCAAGGGGCTCTTACGGGTGATATACCC 3454
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyGluValGlnIleValSerThra 40
DB AGCCTAACCTGCGCGGCAAAAACCAAGTGAAGGTAGGTCCAGATTGTCAACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB GCCCAAACTTCTCCGGCAAGTGCATCATGGGTGTCTGACTGTCTACACAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProIyGlyProValIleGlnMetTyrThrAsnValAsp 80
DB GGAACGAGGACCATCGGTCACCCCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 3634
QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB CAAGACCTTGTGGCTGGCCCGCGCCGCAAGTAGGCGCTCATTTAGACACCCCTGACTTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB GGCCTCTCCGACCTTTACCTGTGTACAGAGGACGCCCATGTCTATCCCGTGGCCGGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
DB GGTGATAGCAGGGGCACTGCTGTCTGCCCGGCCCATTTCTTCAATTTGAAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuLeuProAlaGlyHisAlaValGlyIlePheAlaAlaValCys 160
DB GGGGGTCCGCTGTGTGCCCCCGGGGACGCGCTGGGCATATTATAGGCGCGCTGTGC 3874
QY 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB ACCGTGAGTGTGCTTACGGCGGTGACTTATTCCTTGTGAGAACCTTAGAGAACCACTATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB AGGTCCCGGGTGTTCACGATTACTCTCTCCACCACTAGTGGCCCAAGCTTCCAGCTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrIysValProAlaAlaIaIaYrAla 220
DB GCTCACTCTCATGTCTCCACAGGACGGGCAAAAGCAACAGTCCCGGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB GCTCAGGGCTTAAAGGTAGTACTAACCCCTCTGTGTGTGCAACATGGGGTTTGTGT 4114
QY 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB GCTTACATGTCCAAGGCTCATGGGATCATCTTAATCATCAAGACCGGGGTGAGAACATTT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB ACCACTGGACGCCCATCAAGTACTCCACTACGGCAAGTTCCTTGGCAGCGGGGTGTC 4234
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB TCCGGGGGGCCCTTATGATTAATTTTGTGACAGTGCACCTCCACAGATGCCACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlyThrAlaGlyAlaArgLeuThrVal 320

DB ATCTGGGCATTCGCACTGCTCTTGACCAAGACAGACTCGGGGGGAGACTGTTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB CTGGCCACCGCCACCCCTCGGGCTCGGTACTGTGCCCCCATCCCAATCGAGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGGAGGCTATCCCTCGAAGTATTC 4474
QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB AAGGGGGGAGACATCTCATCTTCTGTCAAGAGAAAGTGCAGAACTCGCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleTyrThrArgGlyLeuAspValSerVal 400
DB AAGCTGTGCAATTTGGCATCAATGCGGTGCTTACTACCGCGGTCTTGACGTGTCCGT 4594
QY 401 IleProThrSerGlyAspValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
DB ATCCGACCAAGGGCGCATGTGTGTGTGTGCAACCGATGCCCTCATGACCGGCTAATCC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB GGGCACTTCGACTCGGTGATGAGCTCAATACGTGTGTCAACCCAGACAGTCGATTCAGC 4714
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB CTTGACCTTACCTTCAACATTGAGACATACCTCCCGGAGATGTGTCTCCGGACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB CAACGTGCGGGCAGAGCTGGCAAGGGGGAAGCAAGGATCTACAGATTTGTGGACCGGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB GAGCGCCCTCCCGCATGTGTGCACTGCTGTCTGTGATGATGACGAGGCTGT 4894
QY 501 AlaTrpTrpGluLeuThrProAlaGlnThrThrValArgLeuArgAlaTyrMetLeuThr 520
DB GCTTGTATGAGCTTCACGCCCGCGGAGACTACGTTAAGGTACGAGGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisIleGluPheTrpGluGlyValPheThrGlyLeu 540
DB CCGGGCTTCCCTGTGTGCGACAGACCATCTTGATTTTGGAGGGCGCTTTACAGGCTTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB ACTCATATAGATGCCCACTTCTATCCAGACAAACAGAGTGGGAGAACCTTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB CTGTATGCGTACCAACCAACCGGTGTGCGGTAGGGCTCAAGCCCTCCCATGTGGGAC 5134
QY 581 GluMetTrpLysCysLeuIleArgLeuLysProThrThrLeuHisGlyProThrProLeuLeu 600
DB CAGATGTGAAGTGTGTTGATTCGCTTCAGGCCCACTCATGGGCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB TACAGACTGGCGCTGTTCAGATGAATCAACCTCAGCAGCACCAATGATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB ATGACATGACATGTGTGCGCCAGCTCGAGGTGTGTCACAGACCTGGGTGCTGTGGCGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB GTCTGTGCTCTTGTGGCCGCGTATTCCTGTCAACAGGCTGCTGATAGTGGCAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys 500
DB 4835 GAGGGCCCTCCGCGCATGTTGCAGCTGCTCCGCTCTGTGAGTGTATGACGACGGCTGT 4894
QY 501 AlaTyrTyrGlyLeuThrProAlaGlyThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGGTATGAGCTCAGCGCCCGCGAGACTACAGTTAGCTACAGAGCTACAGAACACC 4954
QY 521 ProGlyLeuProValCysGlyAspHisLeuGlyPheThrGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTCCCGGTGCGCAGACCATCTTGAATTTGGAGGGCGCTTTACAGGGCTC 5014
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 5015 ACTCATATATAGATGCCCTTCTATCCAGACAAAGAGAGTGGGGAGAACCTTCCCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
DB 5075 CTGGTAGCTTACCAAGCCAGCCAGTGTGGCTACAGGCTCAAGCCCTCCCATCGTGGAGC 5134
QY 581 GluMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGAGAGTGTGATTGCTGCTAAGGCCACCTCCATGGGCCCAACACCCCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGlyValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTCAGATGAAATACCTCGACGACCCAGTCAACCAATATATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGlyValValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATGATGATGTGCGCGCAGCTGTGAGGTGCTACAGACCTGGGTGCTGTGGCGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrLysCysValValIleValGlyArg 660
DB 5315 GTCTGTGCTCTTGGCGCGGTATGCTGTCAACAGCGCTGCGTGCATATAGTGGAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlyValLeuTyrArgGlyPhe 680
DB 5375 GTGCTCTTGTCCGGGAGCCGGCATATACCTGACAGGAGAGCTCTTACCAAGAGTTC 5434
QY 681 AspGluMetGluGlyCys 686
DB 5435 GATGATGATGAGAGTGC 5452
RESULT 12
AR166930
LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION AR166930
VERSION AR166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 9379)
AUTHORS Barban, V.
TITLES Fusion polypeptide having the C protein and E1 protein of hepatitis C virus
JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES Location/Qualifiers
source 1..9379
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,45e-200 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR166930 (1-9379)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3395 CTGGCGCCATTCAGCGGGTATGCGCCAGCACAGAGGGGCTCTTAGGGTCAATATCACC 3454
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 3455 AGCTTAACCTAGCGCGGAGCAAAACCAAGTGAAGGTGAGGTCCAGATTGTGTCAACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 3515 GCCCAAACTTCTCTGCAACGTGCATCATATGGGGTGTGGAGCTGTCTTACCAAGGGGCC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3575 GGAAACAGACATCCCGTACCCAAAGGCTCTGTATCCAGATGATATACCAATGTAGAC 3634
QY 81 GluAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3635 CAAGACTTGTGGGCTGGCGCGCTCCGCAAGTAGCGGCTCATTTGACACCTGCACATGCG 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3695 GGCTCTCGGACCTTACCTGATCACAGAGCACCGCATGTCAATCCCGTCCGCGCGG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 3755 GGATATAGCAGGGGAGCGCTGCTGTGCCCCCGCCCAATTTCTTACTTGAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyTyrPheArgAlaAlaValCys 160
DB 3815 GGGGGTCCGCTGTGTGTGCCCCGGGGCAGCGCGTGGCATATTAGGGCGGGGTGTGC 3874
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
DB 3875 ACCCGTAGAGTGTAGGGGGGTGACTTATCTCTGTGAGAGAACTAGAGAACACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3935 AGGTCCCGGTGTTCACAGATATCTCTTCCACAGATATGCCCCAGACTTCCAGAGT 3994
QY 201 AlaHisLeuHisAlaProThrArgLysSerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTCACTCATCTCTCCACAGGCAAGCGGAAAGACCAAGTCCCGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GTCAGGGGCTATAGGTGTAGTACATCAACCTCTGTGTGCTGCAACATCGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4115 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTACATCAGACCGGGGTGAACAACATT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGCAGCCCATCATCATCTCTCACTCAGGCAAGTCTTGTCCACCGCGGGTGC 4234
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGGCGCTTATGATCATATATTTGAGAGAGTGCACATCCACGATGCCATCTCC 4294
QY 301 IleLeuGlyTyrIleGlyThrValLeuAspGlnAlaGlyThrAlaGlyValArgLeuThrVal 320
DB 4295 ATCTTGGCATTCGCACTGTCTTGTACCAAGCAGAGACTCGGGGGGAGACTGTTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGlyVal 340
DB 4355 CTGGCCACCGCCACCTCCGGGCTCCGTACATGTGCTCCCATCCAAATCGAGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyValAlaIleProLeuGluAlaIle 360

Db 4415 GCTCTGTCACACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAACTATC 4474
 QY 361 LygG1yG1yArgh1sleu1lePheCyB1seSer1yLe1ySeCyAspG1u1eu1a1a 380
 Db 4475 AAGGGGGGAGACATTCATCTCTGTCAATCAAAAGAGAGCGACGACGCGCGCA 4534
 QY 381 LyseuVal1a1euG1yVal1a1aVal1a1a1yTyTyrG1yLeuAspVal1SerVal 400
 Db 4535 AAGCTGTTCGATTTGGGCATCAATGCGGTGCTACTACCGCGTCTTGACGTGTCGTC 4594
 QY 401 1leProThrSerG1yAspVal1a1aVal1a1a1eThrAspAla1eumetThrG1yPheThr 420
 Db 4595 ATCCGACCAAGCGGAGTGTGTGTGTGCAACGATGCTCTCATGACCGGCTATAC 4654
 QY 421 G1yAspPheAspSerVal1leAspCyAsnThrCyAsVal1ThrG1ThrVal1AspPheSer 440
 Db 4655 GGCACCTTCGACTCGGTGATGATGACATACGATGTGTACCCAGACAGTGCATTCACG 4714
 QY 441 LeuAspProThrPheThr1leG1uThr1leThr1eThr1eThr1eThr1eThr1eThr 460
 Db 4715 CTTGACCTTACCTTCACCATTCAGACATCAAGCTCCCGAGAGTGTGTCTCCGCGCT 4774
 QY 461 G1nArgArgG1yArgThrG1yArgG1yLeuProG1y1leTyrArgPheVal1a1aProG1y 480
 Db 4775 CAACGTCGGGGGAGACCTGGACAGGGGAGGAGGCAAGGCAATCTACAGATTGTGTGCG 4834
 QY 481 G1uArgProSerG1yMetPheAspSerSerVal1eucCySerG1yCyTyrArgPhe1a1yG1y 500
 Db 4835 GAGCCGCCCTCCGCGCATGTTCGACTCGTCCGCTCTGTGAGTGTGATGACGAGCGCTGT 4894
 QY 501 AlaTPrpTyrG1u1eThrThrPro1a1uThrThrVal1ArgLeuArgAla1yZMetAsnThr 520
 Db 4895 GCTTGTGTAGACTACGCGCCCGCCAGACTACAGTACGATACGAGCTACATGACACACC 4954
 QY 521 ProG1yLeuProVal1CySerG1nAspH1sleuG1uPheTyrG1uG1yVal1PheThrG1yLeu 540
 Db 4955 CCGGGGCTTCCTCGGTGTGCGACGACCATCTTGAAATTTGGAGGGGTCTTTACAGGGCTTC 5014
 QY 541 ThrH1s1leAspAlaH1sPheLeuSerG1nThrLySG1nSerG1yG1uAsnLeuProTyr 560
 Db 5015 ACTCAATAGATGCCCATCTTCTATCCAGACAAAGAGTGGGAGAACTTCCTTAC 5074
 QY 561 LeuVal1a1aTyrG1nAlaThrVal1CysAlaArgAlaG1nAla1aProProProSerTyrAsp 580
 Db 5075 CTGTGTCGTAACCAACCAACCGGTGTGCGTACAGGCTCCCAAGCCCTCCCAATCGTGAGC 5134
 QY 581 G1MetTyrLySCysLeu1leArgLeuLySProThr1eThrH1sG1yProThrProLeuLeu 600
 Db 5135 CAGATGTGGAAAGTGTGTGATTCGCTTCAGAGCCCACTCCATAGGCGCAACCTCGCTA 5194
 QY 601 TyrArgLeuG1yAlaVal1G1nAsnG1uVal1ThrLeuThrH1sProVal1ThrLySTyr1le 620
 Db 5195 TACAGACTGGGCGCTGTTCAGATGAATACACCTTCAGCCACCAAGTCAACAAATACATC 5254
 QY 621 MetThrCyMetSerAlaAspLeuG1uVal1ThrSerThrTyrVal1LeuVal1G1yG1y 640
 Db 5255 ATGACATGATGTGCGCGGACCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5314
 QY 641 Val1eVal1a1a1eVal1a 660
 Db 5315 GTCTGTGCTGT 5374
 QY 661 1leVal1eSerG1yLySProAla1le1leProAspArgG1uVal1eThrArgG1uPhe 680
 Db 5375 GTCTGT 5434
 QY 681 AspG1uMetG1uG1yCys 686
 Db 5435 GATGAGATGGAAGAGTGC 5452
 RESULT 13
 AR301300
 LOCUS AR301300 9379 bp DNA linear PAT 12-JUN-2003

DEFINITION Sequence 1 from patent US 6538123.
 ACCESSION AR301300
 VERSION AR301300.1 GI:31689076
 KEYWORDS
 SOURCE
 ORGANISM
 SOURCE Unknown.
 REFERENCE 1 (bases 1 to 9379)
 AUTHORS Barhan, V.
 TITLE Vaccine composition for preventing or treating hepatitis C
 JOURNAL Patent: US 6538123-A 1 25-MAR-2003;
 FEATURES
 source Location/Qualifiers
 1..9379
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores: 3.45e-200 Length: 9379
 Pred. No.: 3574.00 Matches: 672
 Score: 99.42% Conservative: 10
 Percent Similarity: 97.96% Mismatches: 4
 Best Local Similarity: 98.78% Indels: 0
 Query Match: 6 Gaps: 0
 DB:

US-09-930-591-2 (1-686) x AR301300 (1-9379)
 QY 1 MetAlaPro1leThrAlaTyrAlaG1nG1nThrArgG1yLeuLeuG1yCys1le1leThr 20
 Db 3395 CTTGGGCGCCATCAGCGCGTACCCGACGACGACAGGCGCTCTAGGGTGCATATTCACC 3454
 QY 21 Ser1eThrThrG1yArgAspLySAsnG1nVal1G1uG1yG1uVal1G1n1leVal1SerThrAla 40
 Db 3455 AGCTTAACCTGGCCGGGACAAACCAAGTGAAGGTGATGCTCAAGATGTGTCAACTGCT 3514
 QY 41 AlaG1nThrPheLeuAlaThrCyS1leAsnG1yVal1CySTPrThrVal1TyH1sG1yAla 60
 Db 3515 GCCCAACCTTCCTGTGCAACGTCATCATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3574
 QY 61 G1yThrArgThr1leAla1eSerProLySG1yProVal1leG1MetTyrThrAsnVal1Asp 80
 Db 3575 GGAACAGAGCCATTCGTCACCAAGGTCCTGTTCATCAATGATGATACCAATGTAGC 3634
 QY 81 G1nAspLeuVal1G1yTyrProAlaProG1nG1yAlaArgSerLeuThrProCySThrCys 100
 Db 3635 CAAGACTTGTGGGCTGTGCGCGCTCCGACAGGATGCGCTCATGTACACCTGCACCTGC 3694
 QY 101 GlySerSerAspLeuTyrLeuVal1ThrArgH1sAlaAspVal1leProVal1ArgArgArg 120
 Db 3695 GGCTCCTCGGACCTTTACCTGTGTACAGAGGACAGCGGATGTATTCCTGCGCGCGCGG 3754
 QY 121 G1yAspG1yArgG1ySerLeuLeuSerProArgPro1leSerTyrLeuLySG1ySerSer 140
 Db 3755 GGTATAGCACAGGGGACGCTGT 3814
 QY 141 G1yG1yProLeuLeuCySProAlaG1yH1sAlaVal1G1y1lePheArgAla1a1a1a1a1a1a 160
 Db 3815 GGGGCTCGCGT 3874
 QY 161 ThrArgG1yVal1a 180
 Db 3875 ACCGTGAGT 3934
 QY 181 ArgSerProVal1PheSerAspAsnSerSerProProAlaVal1ProG1nSerTyrG1nVal 200
 Db 3935 AGGTCCCGGT 3994
 QY 201 AlaH1sLeuH1sAlaProThrG1ySerG1yLyS1eThrLySVal1ProAla1a1a1a1a1a1a 220
 Db 3995 GCTCACCTTCATGTCTCCACAGGACGCGCAAAAGCACAAAGTCCGCGCTGCATATGCA 4054
 QY 221 AlaG1nG1yTyrLySVal1eVal1eVal1eVal1eVal1eVal1eVal1eVal1eVal1eVal 240

Db 4055 GCTCAGGGCTAATGAGTCTAGTACCAACCCCTGTGGTCTGCACACCTGGGCTTTGGT 4114
 Qy 241 AATATMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4115 GCTTAATGTCCTCAAGGCTCATGGATGATCTTACATCAGACCGGGGTGAGAACATTT 4174
 Qy 261 ThrThrGlySerProIleThrIleThrIleThrIleThrIleThrIleThrIleThrIle 280
 Db 4175 ACCACTGGCAGCCCATCACTACTCCACTACGGCAAGTTCTTGGCGAGCGGGGTGC 4234
 Qy 281 SerGlyValAlaIleArgThrIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4235 TCGGGGGCGCTTATGATGATTAATTTGTGACGAGTGCATCTCCAGCGATGCGACATCC 4294
 Qy 301 IleGluGlyIleGlyThrValIleAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4295 ATCTTGGGCACTCGGCACTGCTTACCAAGACAGACTCGGGGGCGAAGCTGTTGTG 4354
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4355 CTGCGCACCGCCACCCCTCGGGGTCTGTCATGTCGCCATCCCAATCGAGAGGTT 4414
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyValAlaIleProLeuGluAlaIle 360
 Db 4415 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCGAAGTATTC 4474
 Qy 361 LysGlyValArgHisIleLeuIlePheCysHisSerIleValIleCysAspGluLeuAlaAla 380
 Db 4475 AAGGGGGGAGACATCTCATCTCTGTATTCATCAAGAGAGAGTGCAGACATCGCCGCA 4534
 Qy 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaIleThrIleArgGlyLeuAspValSerVal 400
 Db 4535 AACCTGGTGCATTTGGGCATCAATGCGGTGCTACTACCGCGTCTTGAAGTGTCCGTC 4594
 Qy 401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrIlePheThr 420
 Db 4595 ATCCCGACCGACGGCGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4654
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4655 GGGGACTTCGACTGGGATGAGTGCATGCTGTCTGATGATGATGATGATGATGATGATG 4714
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4715 CTGACCCCTACCTTCAACATTTAGACATCAAGCTGCTCCCGAGATGCTGTCCCGCACT 4774
 Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
 Db 4775 CAACGTGCGGGGAGGATGCGGAGGGGAGGAGGATCTTACAGATTGTGTGCGACCGGGG 4834
 Qy 481 GluArgProSerGlyMetPheAspSerValLeuGlyCysGlyThrAspAlaGlyCys 500
 Db 4835 GAGGGCCCTCCCGGCAATGTCGACTGTCGCTCTGTGATGATGATGATGATGATGATG 4894
 Qy 501 AlaTrpIleGluLeuThrProAlaGluThrThrValArgLeuArgAlaIleMetAsnThr 520
 Db 4895 GCTTGGATAGAGCTCAAGCCCGCGAGACATCAAGTTAGGCTACGAGGCTACATGAACACC 4954
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 4955 CCGGGGCTTCCTCGGTGCTCCAGACCATCTGAATTTGGGAGGCGCTTTTACAGGGCTC 5014
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValIleLeuLeuProTrp 560
 Db 5015 ACTCATTAAGATGCGCATTTCTATCCCAAGCAAGAGAGTGGGAGAACCTTCTTAC 5074
 Qy 561 LeuValAlaIleArgAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 5075 CTGTGATCGTACCAAGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5134
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuIleArgLeuIleArgLeuIleArgLeuIleArgLeu 600
 Db 5135 CAGATGTGAAGTGTGTGATTCGCTCAAGCCACCTCCATGGGCGCAACACCCCTGCTA 5194

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleIle 620
 Db 5195 TACAGACTGGGCGCTCTTCAAGATGAATATCAACCTGACCAACCCATGACCAATATCATC 5254
 Qy 621 MetThrCysMetSerAlaAspLeuGluValIleThrSerThrTrpValLeuValGlyIle 640
 Db 5255 ATGACATGATGATGCGGCCCAACCTCGAGGTGCTGACAGACCTCGGTGCTGCTGCTGCTG 5314
 Qy 641 ValLeuAlaAlaLeuAlaIleAlaIleThrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 5315 GTCTGCTGCTGCTTGGCGCGGCTGATTCCTGTCAACAGGCTGCTGCTGCTGCTGCTGCTG 5374
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuIleArgGluPhe 680
 Db 5375 GTCTGCTGCTGCGGAGAGCGGCAATCATCTGACAGAGAGTCTTACCGAGATTC 5434
 Qy 681 AspGluMetGluGluCys 686
 Db 5435 GATGAGATGAGAGAGTGC 5452

RESULT 14
 AR176483 9401 bp DNA linear PART 17-DEC-2001
 LOCUS AR176483
 DEFINITION Sequence 9 from patent US 6312889.
 ACCESSION AR176483
 VERSION AR176483.1 GI:17918838
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9401)
 AUTHORS Houghton, M., Choo, Q.-L., and Kuo, G.
 TITLE Combinations of hepatitis C virus (HCV) antigens for use in
 JOURNAL immunosays for anti-HCV antibodies
 PATENT: US 6312889-A 9 06-NOV-2001;
 FEATURES
 source 1..9401
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 3..46e-200 Length: 9401
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: Gaps: 0

US-09-930-591-2 (1-686) x AR176483 (1-9401)
 Qy 1 MetAlaProIleThrAlaIleGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3417 CTGGCGCCCATTCACCGGCGTACGCCAGCAAGAGGCGCTCTTACGGTGCATTAATCACC 3476
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3477 AGCTTACCTGGCGGAGCAAAACCAAGTGAAGGATGAGGTCCAGATTGTGTCACTGCT 3536
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleGlyAla 60
 Db 3537 GCCCAACCTTCTTGGCAAGTGCATCAATGGGATGCTGAGCTGTCTACACGGGGCC 3596
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetIleThrAsnValAsp 80
 Db 3597 GGAACGAGGACATCGCGTACCAAGGCTCTGTCAATCCAGATGATACCAATGTAGAC 3656
 Qy 81 GluAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3657 CAGAACCTTGTGGCTGGCGCGCTCCGCAAGTGAAGCGCTCATTTGACACCTGCACTGC 3716
 Qy 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

```

Db      3717 GGGTCTCCGGACCTTTACCTGTCACAGAGCAGCGAGTGCATTCCTCCGTGGCCGGCGG 3776
Qy      121  G1YASPG1YARGLYSERLEULEUSERPROARGPROILESERLYRLEULYSG1YSESR 140
Db      3777 GGTGATGCGAGGGGCGACCTGCTGTGCCCCGCCCAATTCCTACTTGAAGAGCTCTCG 3836
Qy      141  G1G1YPROLEULEUCYSPROIALG1YH1EALAVAG1Y1EPHEARG1AALAVAG1Y 160
Db      3837 GGGGGTCCGCTGTTGTGCCCCGGGGGCAAGCGGTGGGCAATATTAGGCGCGGCTGTGC 3896
Qy      161  THRARG1YVALALAYVALAVALLASPPHE1LEPROVALG1USERLEU1GUTHRTHMET 180
Db      3897 ACCCGTAGAGGTGAGCGGTGAGCTTATCCCTGTGAGAACCTTAGAGACACCAAG 3956
Qy      181  ARGSEPROVALPHESERASPAENSESERSEPROVALAVALLPROG1SERLYRGINVAL 200
Db      3957 AGGTCCCGGGTTTACGAGTAACCTCTCTCCACCACTAGTGGCCCGAGACTTCCAGGTG 4016
Qy      201  ALAHSLEUHI1SALAPROTHRG1YSEGLYLSERTHRYVALPROVALA1YRVALA 220
Db      4017 GCTACACCTCCATGCTCCACAGCAGCGGCAAAAGCACCAAGTCCCGCTGCATATGCA 4076
Qy      221  ALAAGNGLYTYLYSVALLEUVALLEUASNPSESERVALA1A1ATHMETGLYPHEGLY 240
Db      4077 GCTCAGGGCTATYAGGTGCTAGTACTCAACCCCTGCTGTGCAACAACCTGGCTTGTG 4136
Qy      241  ALATYMETSERLYSALAH1SG1Y1EASPPROAN1LEARGTHRG1YVALARGTHR1E 260
Db      4137 GCTTACATGTCCAAAGGCTCATGGAGTGCATCTTAACATCAGGACGGGGTGAGAACAT 4196
Qy      261  THRTHRG1YSEPRO1LETHRYSERTHRYRGLYSPHELEU1AASPG1YGLYCYE 280
Db      4197 ACCATGCGAGCCCATCAGTACGATCCACCTACGCAAGTCTCTGCGAGCGCGGCTGC 4256
Qy      281  SERGLYGLYALATYRASP1LE1ELECYASPG1UCYS1SERTHRASPALATHRSE 300
Db      4257 TCGGGGGGCGCTTATGACATTAATTGTGACAGATGCCACTCAGAGATGCCACATCC 4316
Qy      301  ILELEUG1Y1EG1YTHRVALLEUASPG1NALAG1UTHRALAG1YVALARGLEUTHRVAL 320
Db      4317 ACTTGGGCGATCGGACGCTCTTGACCAAGCAGAGCTGCGGGGCGAGACTGTGTGTG 4376
Qy      321  LEUVALATHRALATHPROBGLYSEVALTHRVALPROH1SPROAN1LEGL1G1UVAL 340
Db      4377 CTGCGCACCGCACCCCTCGGAGCTCGCTACTGTGCCATCCCAACATGAGAGGTT 4436
Qy      341  ALALEUSERTHRG1YGL1U1LEPROPHETRYRGLYLSA1A1EPROLEUG1UVAL1E 360
Db      4437 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAGATATTC 4496
Qy      361  LYSGLYGLYARGH1SLEUL1EPHECYSH1SERLYSVALYSCYASPG1UENU1A1A 380
Db      4497 AAGGGGGGAGACATCTCATCTTCTGTCTTAAGAAGAAATGGGAGCAACTCCGCCGA 4556
Qy      381  LYSLEUVAL1ALEUG1YVALAEN1AVALA1ATYTYRARG1YLEUASPV1SERVAL 400
Db      4557 AACCTGTGCGATTGGGCATCATGCGCGCTACTACCGGGCTTGAACGTGTCGCTC 4616
Qy      401  ILEPROTHSERGLYASPV1VALVALVALA1ATHRASP1ALEUMETHRGLYPHETHR 420
Db      4617 ATCCGACACGCGCGCATGTTGTCTGTGGCAACGAGTCCCTCATGACCGGCTATACC 4676
Qy      421  GLYASPPHEASPSERVAL1LEASP1YASANTHRYVAL1THRGLINTHRVALASPHESE 440
Db      4677 GGGGACTTCGACTGGGATGACCTGCAATACGTGTGTCAACCAAGCAATTCATTTCAAC 4736
Qy      441  LEUASPPROTHRPHETHR1LEGL1UTHR1LETHRLEUPROGLN1ASPV1AVALSERARGTHR 460
Db      4737 CTTGACCTTACCTTACCATTTAGACATCAACGCTCCCGAGATCTGTCTCCCGCACCT 4796
Qy      461  GLNARGARG1YARGTHRG1YARGGLYLSERPROGLY1LETHRARGPHEVAL1APROGLY 480

```

```

Db      4797 CAACGTGGGGCAGAGACTGGCAGGGGAGAACCCAGCATCTTACAGATTGTGACACCGGGG 4856
Qy      481  GLUARGPROSERG1YMETPHEASPSERVALLEUCYSG1UCYSTRYRASP1A1GLYCYE 500
Db      4857 GAGGCCCCCTCCGCAATGTTGACTGCTCCGTCTCTGTAGTGTCTATGACGAGGCTGT 4916
Qy      501  ALATRYRGL1U1LEUTHRPROVALAG1UTHRTHRVALARGLEUARG1A1YMETASANTHR 520
Db      4917 GCTTGATGAGCTCAGCGCCCGGAGACTACAGTTAGGCTACAGGCTACATGAMACACC 4976
Qy      521  PROGLYLEUPROVALCYSG1NASPH1SLEU1G1UPHE1RPG1UG1YVAL1PHE1RGLYLEU 540
Db      4977 CCGGGGCTTCCCGTGTGCAGAGCACTTGAATTTGGAGGGCGCTTTTACAGGCTC 5036
Qy      541  THRHS1LEASPA1AH1EPHELEUSERGLINTHRYS1G1N1SERGLY1U1AENLEUPROTYR 560
Db      5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAAGAGAGTGGAGAACTTCTCTTAC 5096
Qy      561  LEUVALA1ATYRGINALATHRVALCYEAL1ARG1AG1NALAPROPROPROSE1TPASP 580
Db      5097 CTGTAGAGGTACCAAGCACCGTGTGCGCTAGGGCTCAAGCCCTCCCTCCATCGTGGAC 5156
Qy      581  GLNMETTRP1YSCYSELEU1LEARGLEU1YSPROTHRLEUHI1SG1YPROTHRPROLEULEU 600
Db      5157 CAGATGTGAGAGTGTTAATGTGCTCAAGCCCACTCCATGGGCCCAACCCCTGCTA 5216
Qy      601  TYRARGLEUG1YVALA1G1N1ASNG1UVAL1THRLEUTHRHS1PROVAL1THRYS1Y1E 620
Db      5217 TACAGACTGGGCGCTGTCAAGATGAATCAACCCGACGACCCGATCACCMAATATCATC 5276
Qy      621  METTHR1CYMETSER1A1ASPLEUG1UVAL1THRSE1RTPRVALLEUVALG1YGLY 640
Db      5277 ATGACATGATGATGCGCGACCTGAGAGTGTGCAAGACACTGGGTCTGTGGCGGC 5336
Qy      641  VALLEUVALA1ALEU1A1ALATYR1CYSEUSERTHRG1YCYEVAL1LEUVALG1YARG 660
Db      5337 GTCTGTGCTGCTTTGGCCCGCATTTGCTGTCAACAGGCTGCGTCAATGTTGGCGAG 5396
Qy      661  ILEVALLEUSERGLY1YSPROVAL1LE1EPROASPA1RGL1UVALLEUTHRARG1U1PHE 680
Db      5397 GTGCTGTGTCCGGAGAGCCGGCAATCATATCTGACAGGAGATCTCTTACCGAAGATTC 5456
Qy      681  ASPGLUMETGLUG1UCYS 686
Db      5457 GATGAGATGAGAGATGC 5474

RESULT 15
LOCUS      B6593              9401 bp    DNA          linear    PAT 18-UN-2001
DEFINITION Hepatitis C virus asialoglycoprotein.
ACCESSION  B6593
VERSION    B6593.1 GI:13026028
KEYWORDS  JP 1999071395-A/1.
SOURCE    Hepatitis C virus
ORGANISM  Hepatitis C virus
VIRUSES: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
REFERENCE  1 (bases 1 to 9401)
AUTHORS   Robert,O.R., Frank,M., Kent,B.S., Barbara,A.J. and John,A.H.
TITLE     Hepatitis C virus asialoglycoprotein
PATENT    JP 1999071395-A 1 16-MAR-1999;
JOURNAL   CHIRON CORP
OS        Hepatitis C virus
PN        JP 1999071395-A/1
PD        16-MAR-1999
PF        14-APR-1998 JP 1998103178
PR        08-NOV-1990 US 611,965, 09-NOV-1990 US 611,419 PR
PI        13-SEP-1991 US 758,880
PI        ROBERT O BARUSUTON, FRANK MARCUS, KENT B SUDIAMU, PI BARBARA A
PC        JABAST, JOHN A HALL
C07K14/18, A61K39/29, G01N33/576//C12N15/09, C12P21/02, (C12P21/02, PC
C12R1:91),

```


Db	5277	ATGACATGCATGTCGGCCGACCTGGAGGTCTGCACGACACCTGGGTGCTGCTGTCGCGCC	5336
QY	641	VallLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	5337	GTCTGTGCTGCTTGGCCGGATTCCTCTGTCAACAGGCTGCGTGTCAATAGTGGCAGG	5396
QY	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	5397	GTCTCTTGTCCGGGAAGCCGCATCATTACTGACAGGGAAGTCTTACCGAAGTTTC	5456
QY	681	AspGluMetGluGluCys	686
Db	5457	GATGAGATGGAAGATGC	5474

Search completed: September 17, 2004, 10:30:07
 Job time : 7286 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 06:44:00 ; Search time 684 Seconds
(without alignments)
4260.620 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618
Sequence: 1 MAFITVACQTRGLGCIIT.....PAIIPREVLVREFDEMEEC 686

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-Q=/cgrn2_1/USFTO_epool/US0930591/runcat_13092004_164952_680/app_query.fasta_1.839
-DB=N_Geneseq_29Jan04 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US0930591 @CGN 1.1 470 @runat_13092004_164952_680 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*

1: geneseqn19808:*\n2: geneseqn19908:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3618	100.0	2061	6	AAD34500 Hepatitis
2	3618	100.0	2061	6	AAD31767 Hepatitis
3	3618	100.0	2061	9	AAD60868 Hepatitis
4	3574	98.8	7310	1	AAN92106 Combined
5	3574	98.8	7310	1	AAN90336 Composite
6	3574	98.8	7310	2	AAQ98221 Hepatitis
7	3574	98.8	8316	3	AAA75296 cDNA sequ
8	3574	98.8	9133	2	AAZ07656 Nucleotid

9	3574	98.8	9185	2	AAQ10566
10	3574	98.8	9185	3	AAA75297
11	3574	98.8	9401	2	AAT12710
12	3574	98.8	9401	2	AAT99981
13	3574	98.8	9401	2	AAV09989
14	3574	98.8	9401	2	AAD35043
15	3571	98.7	6905	1	AAN92103
16	3571	98.7	9185	2	AAN26737
17	3570	98.6	9185	2	AAO05956
18	3568	98.6	6239	4	AAK93669
19	3568	98.6	9185	2	AAK00459
20	3567	98.6	8316	2	AAQ05955
21	3565	98.5	2058	6	ABK15344
22	3565	98.5	2058	6	AAD29795
23	3565	98.5	2058	7	ABX14410
24	3565	98.5	2058	9	ADC06768
25	3565	98.5	9400	2	AAQ21744
26	3555	98.3	5360	1	AAN90327
27	3552	98.2	5300	1	AAN92097
28	3551	98.1	9646	2	AAV59361
29	3551	98.1	9646	6	ABK87285
30	3551	98.1	9646	7	ACA62466
31	3551	98.1	12980	2	AAV59364
32	3551	98.1	12980	6	ABK87286
33	3551	98.1	12980	7	ACA62469
34	3551	98.1	16622	3	AAZ36212
35	3550	98.1	9502	2	AAQ74770
36	3544	98.0	9518	5	AAD03778
37	3544	98.0	9539	2	AAK24833
38	3544	98.0	9539	2	AAK24832
39	3544	98.0	9539	4	AAC86938
40	3544	98.0	9611	5	AAC86646
41	3544	98.0	9611	5	AAC86645
42	3544	98.0	9611	5	AAC86647
43	3544	98.0	9611	5	AAC86648
44	3544	98.0	10803	7	ABX10617
45	3543	97.9	9379	2	AAQ36209

ALIGNMENTS

RESULT 1	
ID	AAD34500
XX	AAD34500 standard; DNA; 2061 BP.
XX	
AC	AAD34500;
XX	
DT	16-JUL-2002 (first entry)
XX	
DE	Hepatitis C virus NS3/4A protein encoding DNA.
XX	
KW	Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW	virucide; gene; de.
XX	
OS	Hepatitis C virus.
XX	
FT	Key
FT	Location/Qualifiers
FT	1..2061
FT	/*tag= a
FT	/product= "HCV NS3/4A protein"
XX	
PD	WO200214362-A2.
XX	
PD	21-FEB-2002.
XX	
PF	15-AUG-2001; 2001WO-IB001774.
XX	
PR	17-AUG-2000; 2000US-0225767P.
PR	29-AUG-2000; 2000US-0229175P.
PR	03-NOV-2000; 2000US-00705547.
XX	
PA	(TRIP-) TRIP AB..

AAQ10566	Hepatitis
AAQ75297	Sense str
AAI27110	Hepatitis
AAT99981	HCV polyp
AAV09989	HCV polyp
AAD35043	Hepatitis
AAN92103	Combined
AAK93669	Sense str
AAQ05956	Sense str
AAK93669	HCV NS3/4A
AAK00459	Hepatitis
AAQ05955	Hepatitis
ABK15344	Hepatitis
AAD29795	HCV-1 NS3
ABX14410	DNA encod
ADC06768	HCV mutan
AAQ21744	Compiled
AAN90327	Hepatitis
AAN92097	Hepatitis
AAV59361	Hepatitis
ABK87285	cDNA enco
ACA62466	HCV H77 C
AAV59364	Hepatitis
ABK87286	Hepatitis
ACA62469	DNA encod
AAZ36212	Nucleotid
AAQ74770	Hepatitis
AAD03778	Hepatitis
AAK24833	Infectiou
AAK24832	Infectiou
AAC86938	Nucleotid
AAC86646	Nucleotid
AAC86645	Nucleotid
AAC86647	Nucleotid
AAC86648	Nucleotid
ABX10617	MKO-Z nuc
AAQ36209	Composite

PI Salberg M;
XX WPI; 2002-339446/37.
DR P-PSDB; AAE21837.
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 1; Page 64-65; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is a DNA encoding HCV NS3/4A protein.
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,44e-245 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-930-591-2 (1-686) x AAD34500 (1-2061)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysAlaIleThr 20
DB 1 ATGGCGCTTATCAAGCGCTATGCCAGACAGCAAGGGCCCTTTGGATGATATATACCC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGlyGluValGlnIleValSerThrAla 40
DB 61 AGCTTGAACCGGCGCGGACAAACCAAGGTGAGGGTGAAGTTCAGATCGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysAlaAsnGlyValCysTrpThrValIleGlyAla 60
DB 121 GCCCAGACTTCTTGGCACTGCACTTACCGGGGAGTGTGACGTCGTCTACCATGAGACC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAACAGAGACCATTTGGCTCACTTAAGGCTCTGTATCCAGATGACACCAATGTGGAC 240
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnValAspSerLeuThrProCysThrCys 100
DB 241 CAAGACCTCGTAGGCTGGCCCGCTCCCAAGGTGCCCGCTCATTAACCAATGACATTGC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCTCTGACCTTTTACCTGTGACGAGGACCGCCATGTCATTTCTGTGCGCCGACGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTATGACAGGGGAGGAGCTGCTTCCGCCCGGCTATCTCTACTTAAAGGCTCCCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheAspAlaAlaValCys 160
DB 421 GGAGGCCCTCTGCTGTGCCCCGACAGACATGCCGTAGGCAATTCAGAGCGCGGTATGC 480
QY 161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180
DB 481 ACCCTGAGAGTGGCTTAAGGCGGTGACTTCACTCCCGTAGAGAGTTTGAACACCACTG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCCGGGTGTCTCAGACAACTCTCCCAACAGCAGTGCCTCCAGAGCTACCAAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaIleVal 220
DB 601 GCCCACTGACGTCTCCACCGGACGGGTAGAGACCAACAGTCCCGCGCATATGCA 660

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTTACAAAGTGTGGTGTCTCAACCCCTCGCTTGCGCAACATGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAAAGGCCCATGAGATTTGATCTTAAATCATGAGATCGGGGTGAGACAAAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 781 ACTACTGGAGCCCGATCAAGATTCACCTTACCGGAAAGTTCTTCCACAGCGGGGTGT 840
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCGAGGGGTGCTTATACATATATATTTGAGCAGATGCCACTCCACGATGCAACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValaArgLeuThrVal 320
DB 901 ATCTTGGGCAATTGGCACTGTCTTGAACAGCAGAGACCGGGGGGAGACTGACTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
DB 961 CTGCGCAACGCTTACCCCTCGGGCTCCGTCACGTGCTCCCATCTTAACATCAGAGAGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGCCACTACCGGAGAGATCCCTTTATGACAGGCTATTCCTTGAAAGCAATT 1080
QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerLysLysGlyCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCACTTCTGCACTCAAAAGAGAGGACAGCGCTGGCGGCA 1140
QY 381 LysLeuValAlaLeuGlyValaAsnAlaValaIleTyrThrArgGlyLeuAspValSerVal 400
DB 1141 AAATCTGTGCGCTTGGGCGTCAATGCCGTGCTTACTACCGGGCTTGAATGTCTCGTC 1200
QY 401 IleProThrSerGlyAspValValaValaIleThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCGACAGATGGTGCCTTGTGCTGTGGCAACAGCGCCCTCATGACCGGCTTTTACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValaThrGlnThrValaAspPheSer 440
DB 1261 GGGGACTTCAATTCGGATGATGACTGCAACAGTGTGTACCCAGACAGTCTGACTTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCAACATTTGAGACATACAGCTTCCCAAGATGTGTCTCCGTAAT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTGGGGTGAAGCTGGCAGAGGGAGCCAGGCAATCAAGATTTGTGGACCGGGG 1440
QY 481 GlnAspProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGGCTCTTCTGCAATGTTGACTGTCTGTCTCTGCAAGTGTCTATGACGGGGTGT 1500
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValaArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTGATAGCTTACGCCCGCGGACAGACAGTATAGCTTACGAGCATATGAAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluValPheThrGlyLeu 540
DB 1561 CCGGACCTTCCCGTGTCCAAAGCAACATCTTGAATTTTGGAGGGCGCTTTTACGGGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACCCAATTAAGCGCCACTTCTATCCCAAGCAAGAGAGTGGGAAACCTTCCCTAT 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 1681 CTGTATACGTCACAAAGCAACGTGTGCGTAAAGCTCAACCCCTCCCGTGTGGGAC 1740

QY	581	GlnMetT ¹ PLysCysLeuIleI ² earGLeuLysP ³ roThrLeuH ¹ IG ¹ lyProThrProLeuLeu	600
Db	1741	CAGAGTGGAAAGTGGTGTGATCCGTCTCAAGCCCACTCTCCAGGGCCAAACACTCTGCTA	1800
QY	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	1801	TATAGACTGGGGCGTGTCCAGAAATGAAGTCAACCTTCAGCCACCCAGTACCAAGTATATC	1860
QY	621	MetThrCysMetSerHisAspLeuGluValValThrSerThrTyrValLeuValGlyGly	640
Db	1861	ATGACATATATAGTCCGCTGACCTGGAGGTCGTACAGAGTACCTGGGTGCTCTGGCGGC	1920
QY	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	1921	GTTCTGGCGTCTTGTGGCCCGCATTTGCTCATTCACAGCGCTGGCTGCTCATAGTAGAGC	1980
QY	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	1981	ATTGCTTGTCTCGGAAAGCCGGCATTATCATACCGACAGGGAAGTCTCTACCGGGAGTTC	2040
QY	681	AspGluMetGluGluCys	686
Db	2041	GATGAATGGAAGAGTGC	2058
RESULT 2			
ID	AA031767	standard; DNA; 2061 BP.	
XX	AA031767;		
XX	AA031767;		
DT	18-JUN-2002.	(first entry)	
XX	Hepatitis C virus (HCV) NS3/4A DNA coding region.		
DE	Hepatitis C virus (HCV) NS3/4A DNA coding region.		
XX	Hepatitis C virus; HCV infection; viraemia; fungicide; antibacterial;		
KW	Hepatitis C virus; HCV infection; viraemia; fungicide; antibacterial;		
KW	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;		
KW	ds.		
XX	Hepatitis C virus.		
OS	Hepatitis C virus.		
XX	Key	Location/Qualifiers	
FH	CDS	1..2061	
FT		/*tag= a	
FT		/product= "HCV NS3/4A protein"	
XX	MO300213855-A2.		
PN	21-FEB-2002.		
XX	21-FEB-2002.		
PD	15-AUG-2001; 2001WO-1B001808.		
XX	15-AUG-2001; 2001WO-1B001808.		
PF	17-AUG-2000; 2000US-0225767P.		
PR	29-AUG-2000; 2000US-0229175P.		
PR	03-NOV-2000; 2000US-00705547.		
PA	(TRIP-) TRIP AB.		
XX	Sallberg M, Hultgren C;		
XX	WPI: 2002-241837/29.		
DR	P-FSDb; AAE19900.		
XX	Vaccine compositions for treating and preventing disease, preferably		
PT	hepatitis C virus infection, comprises ribavirin and antigen that has		
PT	epitope present in hepatitis C virus.		
XX	Claim 1; Page 94-95; 120pp; English.		
XX	The invention relates to a composition comprising ribavirin and an		
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of		
CC	hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV		
CC	sequence. The composition is useful for enhancing an immune response to a		
CC	hepatitis C antigen in humans, domestic, sport or pet species and as		

[illegible]

```

Db      841 TCAGGGGGTGTATGACATATATTGTGACAGGCACTCCAGGATGCAACATCC
Qy      301 ILELENGLYILEGLYTHRAILEUASGLNALAGLTHRAAGLYAILEUAGLYTHRA
Db      901 ATCTGGGCATGTGGACATGCTTGACCAAGACACCGGGGGGAGACGACGATG
Qy      321 LEUATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      961 CTGCCACCGCTACCTCCGGGCTCCGCTACCTGCTCCATCTTAACATGAGAGATT
Qy      341 ALAUSERTHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1021 GCTCTGTCCACTACGAGAGATCCCTTTATGGCAAGGCTATCCCTTGAAGCAATT
Qy      361 LYSGLYGLYTHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1081 AAGGGGGGAGACATCTCATCTCTGCACTCAAGAGAGAGAGAGAGAGAGAGAG
Qy      381 LYSLEUVALALALEUGLYVALASNALVALALATYTRARGGLYLEUASPVALSERVAL
Db      1141 AAATGCTGGCTGGGGGCTGAATGCGGTGCTTACACCGCGCTTGATGTCGCTC
Qy      401 ILEPROTHRSERGLYASPVVALVALVALATHRAAPALAEUMECTHRAATHRA
Db      1201 ATCCCGACAGTGGTGACGTGTGCTGCGCACTGACCGCTCATGACCGCGCTTAC
Qy      421 GLYASPHASPSERVALILEASPCYASATHRAATHRAATHRAATHRAATHRAATHRA
Db      1261 GGCACATTCGATGCGGATGAGTACGACACGAGTGTCCACCAAGAGAGAGAGAGAG
Qy      441 LEUASPROTHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1321 CTGACCTCTACCTTACCATGAGACATACCGCTTCCCGAGATGCTCTCCGACT
Qy      461 GLNARGARGLYARGTHRAARGLYLYSPROGLYILEYTRARGPHEVALALAPROGLY
Db      1381 CAACGTGGGGGTAGAGACTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Qy      481 GLNARGPROSERGLYMETPHEASPSERVALILEUCYSGLYCYTRASPALAGLYCY
Db      1441 GAGGCTCTCTTGGGCAATGTTGACTGCTCTCTCTGCGAGTGTAGACCGGGTGT
Qy      501 ALATPYRYGLIULEUTHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1501 GCTTGGATAGCTTACCGCCGAGACACAGTATGAGTACGAGCATATGAAACCC
Qy      521 PROGLYLEUPROVALCYSGINASPHISLEUGLUBHETPROGLYVALPHEATHRA
Db      1561 CCGGAGACTTCCCGTGTGCAAGACATCTGAAATTTGGGAGGCGCTTTACCGGCTC
Qy      541 THRISILEAPALATHRAATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1621 ACCACATAGACGCCCACTCTCTATCCAGCAAGAGAGAGAGAGAGAGAGAGAG
Qy      561 LEUVALALATYRYGINALATHRAATHRAATHRAATHRAATHRAATHRAATHRA
Db      1681 CTGGATAGGATCAAGACCGGTGTGCTGAGACTCAAGCCCTCCCGCTGAGGAC
Qy      581 GLINMETPYRYSGYLEUILLARGLEULYSPROTHRAATHRAATHRAATHRA
Db      1741 CAGATGTGAGAGTGTGATCCGTCTCAAGCCCACTCCCAAGGAGAGAGAGAG
Qy      601 TYTRARGLEUGLYALALAVALASGLVALATHRAATHRAATHRAATHRAATHRA
Db      1801 TATGACCTGGGCGCTGTCCAGAAATGAGTCACTGAGCAGCAGCAGTCAACAAATATC
Qy      621 METTRCYMETSERIALASPLLEUGLYVALATHRAATHRAATHRAATHRAATHRA
Db      1861 ATGACATGATATGTGGCTGAGCTGAGAGGTGCTCAAGAGTACCTGGGCTGTTGGCGG
Qy      641 VALLEUVALALALEUVALALATYCYSEUSERTHRAATHRAATHRAATHRAATHRA

```

```

Db      1921 GTTCTGCGCTGTTGGCCGGATGTCCTATCCACAGGCTGCGTGTATAGTAGTAGG
Qy      661 ILEVALLEUSERGLYLYSPROVALILEILEPROASPARGLIUVALLIETHRAATHRA
Db      1981 ATTGTCTTGTCCGGAAGCCGCAATCATACCGAGAGAGAGAGAGAGAGAGAG
Qy      681 ASPGLUMETGLIUCYS 686
Db      2041 GATGAATGGAAGAGTGC 2058

RESULT 3
AAB0868
ID AAB0868 standard; DNA; 2061 BP.
XX
AC AAB0868;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A DNA.
XX
KM Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide; de.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 1..2061
FT /tag= a
FT /product= "Hepatitis C virus protein"
XX
FN US2002136740-A1.
XX
XX 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX
XX (SALT/) SALTBERG M.
XX (HULT/) HULTGREN C.
XX
PI Salberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
DR P-PSDB; ABW00351.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 1; Page 60-61; 0pp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis C virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A DNA
XX
SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,44e-245 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-930-591-2 (1-686) x AAB0868 (1-2061)

```

QY 1 MetAlaProIleThraIaIaIaGlnGlnThraArgGlyLeuLeuGlyCysIleIleThr 20
 Db 1 ATGGCGGCTTATCAAGGCTTATGCGCAGACAAAGGGGCTTTGGAGTCAATATACACC 60
 QY 21 SerLeuThrGlyIaArgApIySaAsnGlnValGluGluValGlnIleValSerThrAla 40
 Db 61 AGCTTGACCGGCGGAGCAAAACACAGTGAAGGTTCAGATCGTGTCAACTGCT 120
 QY 41 AlAGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIYThIaGlyAla 60
 Db 121 GCCAGACTTCTTGGACACTGCATTACGGGGTGTGGACTGTCTACATGAGACC 180
 QY 61 GlyThrArgThrIleAlaSerProIyGlyProValIleGlnMetYTrhAsnValAsp 80
 Db 181 GGAACAAGACACATTTGGTCACTTAAGGTCCTGTATCCAGATGTACCAATGTGTGAC 240
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyValArgSerLeuThrProCysThrCys 100
 Db 241 CAAGACCTCGTAGGCTGGCCGCTCCCAAGTGCCTCATTTAACACATGCACTTGC 300
 QY 101 GlySerSerAspLeuYTrLeuValThraArgHisAlaAspValIleProValArgArgArg 120
 Db 301 GGCTCTCGACCTTTACCTGGTCAACAGGACGCGCATGTCTTCTGTGCGCCGACGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYTrLeuGlyGlySerSer 140
 Db 361 GGATATGGACAGGGGCACTGTCTTCCCGCGCTATCTTACTTGAAGAGGCTCTCG 420
 QY 141 GlyGlyProIleuLeuCysProAlaGlyHisAlaValGlyIlePheAlaGlyAlaValCys 160
 Db 421 GGAGGCTCTGTGTGCTCCCGCAGACATGCCGTAGGCATATTCAAGCCGCGGTATGC 480
 QY 161 ThrArgGlyValAlaIySaIaValaIaAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 481 ACCCGTGAAGTGGCTTAAGCGGTGACTTATCCCGTGAAGACTTATGAGACACATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYTrGlnVal 200
 Db 541 AGGTCCCCGGGTCTCAGACAACTCTCCCAACACAGATGCCCAAGACTACCAAGT 600
 QY 201 AlaHisIleuHisAlaProThrGlySerGlyIySerThrIySValProAlaAlaIyAla 220
 Db 601 GCCCACTGCATGTCTCCACCGGACGGGTAAAGACCAAGGTCCTGGCGGCACTAGCA 660
 QY 221 AlAGlnGlyYTrIySValIleuValIleuAsnProSerValAlaIaThrMetGlyPheGly 240
 Db 661 GCTCAGGCTCAAAAGGTGTGTGCTCAACCCCTGCTGTGCAACAATGGGCTTGTGT 720
 QY 241 AlaIyMetSerIySaIaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 721 GCTTACATGTCCAAAGGCCATGGGATTGATCTTAACATCAAGACTGGGGTGAAGCAATT 780
 QY 261 ThrThrGlySerProIleThrYTrSerThrYTrGlyIySValPheLeuAlaAspGlyIyCys 280
 Db 781 ACTACTGGCAGCCGATCAGGTATTCACCTACCGCAAGGTTCTTGGCGACGGGGGTGT 840
 QY 281 SerGlyIyAlaIyTrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 841 TCGAGGGGTGTATGACATATATTTGTGACAGATGCCACTCAGAGATGCAACATCC 900
 QY 301 IleLeuGlyIleGlyThrValIleuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 901 ATCTTGGGCAATTGGCACTGTCTTGACCAAGCAAGACCGGGGGGAGACTGACTGTG 960
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 961 CTCGCGACCGCTACCTCCGCGGCTCCGCTCACTGTGCCCATCTTAACATCGAGAGAGTT 1020
 QY 341 AlaLeuSerThrThrGlyIyGluIleProPheYTrGlyIySaIaIleProLeuGluAlaIle 360
 Db 1021 GCTGTGTCTACCTACCGAGAGATCCCTTTATGACAAAGGCTTATCCCTTGAAGCAATT 1080
 QY 361 LySGlyIyArgHisIleuIlePheCysHisSerIySValCysAspGluLeuAlaIa 380

Db 1081 AAGGGGGGAGACATTCATCTTGTGCCACTCAAAAGAGAGGACAGACTGCGCGCA 1140
 QY 381 LyLeuValAlaLeuGlyIyValAsnAlaValAlaIyTrArgGlyLeuAspValSerVal 400
 Db 1141 AACTGTGGCGCTTGGCGCTCAATGCCGTGGCTTACTTACCGCGCTTATGTGTCCGT 1200
 QY 401 IleProThrSerGlyIyAspValValValAlaIaThrAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACCAAGTGTGACGTGTGCTGTGGCACTGACGCGCTTACGACCGGCTTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 1261 GGGCACTTGATTCGGTGTATGACTCAACACGTGTGTACCCAGACAGTCACTTACG 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACTTCACTTGAACATACACGCTTCCCAAGATCTGTCTCCGTAAT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyIyProGlyIleYTrArgPheValAlaProGly 480
 Db 1381 CAACGTGGGGTATGACTGGCAGAGGAGGAGCAGCATCTACAGATTTGTGGACCGGG 1440
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTrpAspAlaGlyCys 500
 Db 1441 GACGCTCTTCTGGCATGTTTGACTGTCTGTCTGTGAGAGTATACCGGGGTGT 1500
 QY 501 AlaTrpYTrGluLeuThrProAlaGluThrThrValArgLeuArgAlaIyTrMetAsnThr 520
 Db 1501 GCTTGTATGAGCTTACGCCCGCCGACACACAGTTAAGCTACGACATACATGAACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisIleuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 1561 CCGGACTTCCCGTGTGCCAAGACCATCTTGAATTTTGGAGGGCGCTTTACGGGCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIySValGlnSerGlyIyLeuLeuProIy 560
 Db 1621 ACCCACTAGACCCCACTTCTATCCAGACAAACAGAGTGGGAAAACTTCCCTAT 1680
 QY 561 LeuValAlaIyTrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 Db 1681 CTGTGTAGCGTACCAAGCCACCGTGTGCTAGAGTCAAGCCCTTCCCGTGTGGAC 1740
 QY 581 GlnMetTrpIySValLeuIleArgLeuIySProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGAAGTCTTGAATCCGTCTCAAGCCACCTCATGTGGCAACCTCTGCTA 1800
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrIySVal 620
 Db 1801 TATAGACTGGCGCTGTCCAGATGAAGTCAACCTGACCAACCAAGTATATC 1860
 QY 621 MetThrCysMetSerIleAspLeuGluValValThrSerThrTrpValIleuValGlyIy 640
 Db 1861 ATGACATGATATGTGGCTGACCTGAGGTGTGACAGATCTGGGGTGTGTTGGCGGC 1920
 QY 641 ValLeuAlaAlaLeuAlaAlaIyTrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 1921 GTTCTGGCTCTTGTGGCCGGTATTCCTATCAACAGGCTGCGTGTCAATGATGAAGT 1980
 QY 661 IleValLeuSerGlyIySProAlaIleIleProAspArgGluValLeuYTrArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGGAAGCCGCGCAATCATVACCGACAGGAGATCTCTACCGGAGATT 2040
 QY 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGATGC 2058

RESULT 4

AAN92106 standard; DNA; 7310 BP.
 AAN92106;

DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones K9-1 through 15e.
 XX
 KM Hepatitis C virus; HCV, non-A, non-B hepatitis; NANBH.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..7310
 FT /*tag= a
 XX
 XX EP18216-A.
 XX
 PD 31-MAY-1989.
 XX
 PF 18-NOV-1988; 88EP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 PA (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo QL, Kuo G;
 XX
 DR WPI, 1989-159274/22.
 DR P-PSDB; AAP92050.
 XX
 PT Purified hepatitis C virus - and associated nucleic acids and
 PT peptide(s).
 XX
 PS Claim 3; Fig 47-1 - 47-8; 139pp; English.
 XX
 CC It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC
 SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
 XX
 Alignment Scores:
 Pred. No.: 1 42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-930-591-2 (1-686) x AAN92106 (1-7310)
 QY 1 MetalapioleThraAlaTyraAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1728 CTGGGGCCCATCAGCGCGTACGCCAGCAGACAAAGGGGCTCTAGGGGTGCAATATCAC 1787
 QY 21 SerLeuThGlyAArgAspLysAsnGlnValGlnGlyValGlnIleValSerThrAla 40
 DB 1788 AACCTAACTGGCGGAGCAAAAACCAAGTGAAGGTGAGGTCCAGATTGTGTCAACTGCT 1847
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrrPThrValTyrlsGlyAla 60
 DB 1848 GCCCAAACTTCTCTGGCAAGTGCATCATGGGGTGTGTGTGACACTGTCTACACAGGGGCC 1907

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyThrAsnValAsp 80
 DB 1908 GGAACAGAGACCATCGCGTACACCCAAAGGTCCTGTCACTCAGATATATACCAATGTACAC 1967
 QY 81 GlnAspLeuValGlyTrrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1968 CAAGACCTTGTGGTGGCGCCGCTCCGCAAGGTACCGCTCATTTACACCTCGACTTGC 2027
 QY 101 GlySerSerAspLeuTyrlleuValThrArgHisAlaAspValIleProValArgArg 120
 DB 2028 GGCTCTCGGACCTTTACCTGTGTACAGAGCAGCGCGATGTCACTCCGGTGGCGCGCG 2087
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaGProIleSerTyrlleuGlySerSer 140
 DB 2088 GGTGATAGCAGGGGAGCGCTGTCTGTCCCGCCCATTTCTTACTTGAAGGCTTCCTG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 2148 GGGGGTCCCTGTGTGTGCCCCCGGGGACGCGGTGGCATTTTAGCGCGGGTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlySerLeuGlnThrThrMet 180
 DB 2208 ACCCGTGAAGTGGCTAAGCGGAGCTTATCCGTGTGAGAACCTAGACCAACATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrlleuVal 200
 DB 2268 AGGTCCCGGTGTGTACGATTAATCTCTCCACCAAGTGTCCCGAGCTTCAGGTG 2327
 QY 201 AlaHisLeuHisAlaProThrArgLysSerGlyLysSerThrLysValProAlaAlaTyrla 220
 DB 2328 GCTCACTCTCAATCTCCACAGGACGGGCAAAAGCAAGATCCCGCTGCATATGCA 2387
 QY 221 AlaGlnGlyTyrlleuValleuValleuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 2388 GCTCAGGCTATATAGGTGTAGTACCAACCCCTGTGTGTGCAACACTGGGCTTTGGT 2447
 QY 241 AlaTyrlMetSerTyrlleuValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 2448 GCTTACATGTCCAAAGCTCATGGATGATCTTACATACAGACCGGGGTGAGAACAT 2507
 QY 261 ThrThrGlySerProIleThrTyrlSerThrTyrlGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 2508 ACCACTGGACGCCCATCATGATCTCCATCAGGCAAGATTCCTTCCACCGCGGGTGC 2567
 QY 281 SerGlyGlyAlaTyrlAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
 DB 2568 TCCGGGGGCGCTTATGACATTAATTTGTGACAGATGCCACTCCAGGATGCCAATCC 2627
 QY 301 IleLeuGlyTlleGlyThrValIleuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 2628 ATCTTGGGCAATCGGCACTGCTTTCACCAAGACAGACTCGGGGCGAGACTGTGTG 2687
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGluVal 340
 DB 2688 CTGGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCCAAATCGAGAGGT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrlGlyValAlaIleProLeuGlnAlaIle 360
 DB 2748 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGAAGGCTATCCCTTCGAAGTATC 2807
 QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysIleLysCysAspGlnLeuAlaIle 380
 DB 2808 AAGGGGGGAGACATCTCATCTTGTGATTAAGAAAGAGAGCGCAACATCGCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrlTyrlArgGlyLeuAspValSerVal 400
 DB 2868 AAGCTGTGCGCATTTGGCATCAATGCGGTGGCTTACTACCGCGGTCTTGAAGTGTGCTGC 2927
 QY 401 IleProThrSerGlyAspProValIleValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
 DB 2928 ATCCGACCAAGCGCGATGTGTGTGTGTGTGACCAAGATCCCTCATGACCGGCTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

```

Db      2988 GGGGACTTCACTCGGATAGACTGCAATACTGTCTACCCAGACAGTCGATTTCAGC
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr
Db      3048 CTTGACCTTACCTTACCACTTGAACATACGCTCCCGGAGATGCTGTCTCCGCACT
Qy      461 GlnArgArgIleArgThrArgIleArgIleArgProGlyIleThrArgPheValAlaProGly
Db      3108 CAACGTGGGGGAGGAGTGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
Qy      481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyCysIleArgAspAlaGlyCys
Db      3168 GAGGCGCCCTCCGCACTGTCGACTCGCTCCCTGTGAGTGTGAGCGCAGGCGCTGT
Qy      501 AlaTrpTrpGluLeuThrProAlaGluThrValArgLeuArgAlaTrpMetAspThr
Db      3228 GCTTGGTATAGCTCAGCGCCGCGGAGACTACAGTACAGTACAGGCGTACATGACACC
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu
Db      3288 CCGGGGCTTCCCGTGTCCAGGACCATCTTGAATTTGGGAGGGCGCTTTACAGGCTC
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGluAsnLeuProTrp
Db      3348 ACTCATATAGATGCCACTTCTTATCCAGCAAGACAGAGTGGGAGAACTTCTTAC
Qy      561 LeuValAlaTrpGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerTrpAsp
Db      3408 CTGGTACGTTACCAAGCCAGCGTGTGCTAGGGCTCAAGCCCTCCCACTCGTGAC
Qy      581 GlnMetTrpIleCysLeuIleArgLeuLeuProThrLeuHisGlyProThrProLeuLeu
Db      3468 CAGATGTGAAAGTCTTGATTCGCTCAAGCCCACTTCATGGGCGCAACCCCTCTCA
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleTyrIle
Db      3528 TACGACTGGGCGCGCTGTTCAAGATGAATACACCTGACGACCAAGTCAACCAATACATC
Qy      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyIle
Db      3588 ATGACATGATGTCGCGGAGCTGAGAGTGTCTCAAGACCTGGGAGTGTCTGTTGGCGGC
Qy      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg
Db      3648 GTCTGCTGCTGTTGGCGCGGTATTCCTGTCAACAGGCTGCGTGTGATATGAGGCAAG
Qy      661 IleValLeuSerGlyIleProAlaIleIleProAspArgGluValLeuTyrArgGluPhe
Db      3708 GTGCTGTTGTCCGGGAGCGGCAATCATCTGACAGGGAAGTCTTACCGAGAGTTC
Qy      681 AspGluMetGluGluCys 686
Db      3768 GATGATGATGAAAGATGC 3785

```

RESULT 5
 ID AAN90336 standard; DNA, 7310 BP.
 XX AAN90336;
 AC AAN90336;
 DT 25-MAR-2003 (revised)
 DT 19-JUL-2001 (revised)
 DT 01-NOV-1989 (first entry)

XX Composite hepatitis C virus (HCV) cDNA.
 KM Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
 XX Pan troglodytes.
 OS
 XX GB2212511-A.
 PN
 XX

```

PD      26-JUL-1989.
XX      18-NOV-1988; 88GB-00027024.
PF      18-NOV-1988; 88GB-00027024.
XX      18-NOV-1987; 87US-00122714.
PR      30-DEC-1987; 87US-00139886.
PR      26-FEB-1988; 88US-00161072.
PR      26-OCT-1988; 88US-00263584.
XX      (CHIR ) CHIRON CORP.
PA      Houghton M, Choo QL, Kuo G;
PI      WPI; 1989-215054/30.
DR      P-PSDB; AAP90288.
XX      Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT      polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT      infection.
XX      Disclosure; Fig 47; 30pp; English.
XX      The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
CC      aligning clones K9-1 through 15e in 5'-3' direction. The cDNA encodes
CC      antigens which react with antibodies in patients with non-A non-B
CC      hepatitis (NANBH). The cDNA can be used to design probes, or to
CC      synthesize polypeptides, which are used to diagnose HCV-induced NANBH, to
CC      raise antibodies for immunosassay or treatment, or to produce vaccines.
CC      See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to
CC      correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
CC      field.)
SQ      Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.: 1,42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x AAN90336 (1-7310)

```

Qy      1 MetAlaProIleThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr
Db      1728 CTGGCGCCCATTCACGCGCGTACGCCAGACAAAGGGGCTCTTACGGTGCATTAATCACC
Qy      21 SerLeuThrGlyArgAspIleAsnGluValGluGlyValGlnIleValSerThrAla
Db      1788 AGCTTAATCTGCGCGGAGCAAAACCAAGTGAAGGTGAGGTCCAGATTGTCTCACTGCT
Qy      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleGlyAla
Db      1848 GCCCAAACTTCTCTGCAACAGTGCATCAATGGGGTGTGCTGAGCTGTCTACACAGGGGCC
Qy      61 GlyThrArgThrIleLeuSerProIleGluValIleGlnMetIleThrAsnValAsp
Db      1908 GGAACGAGACCATTCGCTACCCAGAGGCTCTGTATCCAGATGATACCAATGATGAC
Qy      81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys
Db      1968 CAAAGACTTGTGGGCTGGCCGCTCCGCAAGTACGCTCATATGACACCTGCACTTGC
Qy      101 GlySerSerAspLeuTrpLeuValThrArgHisAlaAspValIleProValArgArgArg
Db      2028 GGCTCTCGGAGCTTACCTTACCTGTCACAGGACAGCGAGTCAATCCCGGCGCGCGG
Qy      121 GlyAspGlyArgGlySerLeuLeuSerProAlaGlyProIleSerTrpLeuIleGlySerSer
Db      2088 GGTGATGACGAGGGGCGAGCTGTGCTGCGCCCGGCAATTTCTTCACTGAAAGGCTCTCG
Qy      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys

```

```

Db      2148 GGGGGTCCGCTGTTGTGCCCCGGGGGAGCGCGTGCGCATATTATTAAGGGCGCGGTGTC 2207
Qy      161  ThrArgIValAlaValAspPheIleProValGluSerLeuGluThrThMet 180
Db      2208 ACCGTGGAGGTGAAGCGGTGAACCTTATCCCTGGAGAACATGAACAACATG 2267
Qy      181  ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGluVal 200
Db      2268 AGGTCCCGGTGTTTCAAGATTAATCTCTTCCACCAAGTAGGCCCAAGCTTCCAGGTG 2327
Qy      201  AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      2328 GCTCACTCCATGCTCCACAGGAGCGGCAAAAGACCAAGGTCCTCGGCGCATATGCA 2387
Qy      221  AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      2388 GCTAGGGCTAAAGGTGCTAAGTCAACCCCTCTGTGTGCAACATGCGGCTTGTGT 2447
Qy      241  AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      2448 GCTTACATGTCACAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAACAATT 2507
Qy      261  ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      2508 ACCACTGGCAGCCCATCACTACTCCACCTAAGGAACTTCTTGGCCGACGCGGGTGC 2567
Qy      281  SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      2568 TCGGGGGGCGCTTATGACATATATTTGTGACAGAGTCCACTCAGAGATGCCACATCC 2627
Qy      301  IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db      2628 ATCTTGGGCATCGGCACTGCTTGAACAAGAGACTGGGGGGGAGACTGTGTGTG 2687
Qy      321  LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      2688 CTCGCCACCGCACCCCTCCGGGCTCCGTCACGTGCCCCCATCCCAATCGAAGAGTT 2747
Qy      341  AlaIleuSerThrThrGluGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      2748 GCTCTGTCACACCGGAGAGATCCCTTTTACGCAAGGCTATGCCCTGAAATATC 2807
Qy      361  LysGlyGlyArgHisIleLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db      2808 AAGGGGGGAGACATCTCATCTTCTGTATTCAAAGAAAGAGTGGAGAGCTGCCGCA 2867
Qy      381  LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      2868 AAGCTGGTCGATTTGGGATCAATGCCCTGCTTACTACCCGGGTCTTGAGCTGTCCGTC 2927
Qy      401  IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      2928 ATCCGACACGAGGCGATGTTGCTGCTGCGCAACCAATGCCCTCAAGACCGGCTATACC 2987
Qy      421  GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      2988 GCGGACTTCGACTCGGTGATAGACTGCAATAGCTGTGCAACCCAGACAGTGAATTGACG 3047
Qy      441  LeuAspProThrPheThrIleGluThrIleLeuProGlnAspAlaValSerArgThr 460
Db      3048 CTGACCTTACCTTCAACATTAAGACATACGCTCCGCCAGAGTGTCTCCGCACT 3107
Qy      461  GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      3108 CAAGCTCGGGGAGAGACTGCGAGGGGGAGGCAAGCATCTACAGATTTGTGGACCGGGG 3167
Qy      481  GlnArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db      3168 GAGGCGCCCTCCGGGATGTTGACTCGTCCGCTGTGTGAGTGTATGACGAGCGCTGT 3227
Qy      501  AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuAlaGlyAlaTyrMetAsnThr 520

```

```

Db      3228 GCTTGGTATGAGCTACGCCCCGAGACTACAGTTAAGCTACAGCGTACATGAACCC 3287
Qy      521  ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluIleValPheThrGlyLeu 540
Db      3288 CCGGGGCTTCCCGTGTGCGCAGGACATCTTGAAATTTTGGAGGGCGCTTTTACAGGCTC 3347
Qy      541  ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluLeuProTyr 560
Db      3348 ACTCATATATATGCCCATCTTATCTTATCCAGACAAAGCAGAGTGGGAGAACTTCTTAC 3407
Qy      561  LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerPheAsp 580
Db      3408 CTGTATGCGTACCAAGCAGCAGGTGTGCTAGGGCTCAAGCCCTTCCCATGTGGAGAC 3467
Qy      581  GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      3468 CAAGTGTGAAGTGTGATGATGCTTCAAGCCCAACCTCCATGGGCCAACCCCTGTGCTA 3527
Qy      601  TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      3528 TACAGACTGGCGCGCTTTCAGATGAATACACCTGACGACCCAGTCAACCAATATCATC 3587
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      3588 ATGACATGATGATGTCGCGGCACTGAGAGTGTGACAGACCTGTGGTGTCTGTGGCGGC 3647
Qy      641  ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      3648 GTCTGCTGCTCTTGTGGCCCGGTATTCCTGTCAACAGGCTGCTGTGATAGTGGGAGG 3707
Qy      661  IleValIleuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      3708 GTGCTTGTTCGCGGAAAGCCGCAATCATCTGACAGGAGATCCTTACCGAGATTCC 3767
Qy      681  AspGluMetGluGluCys 686
Db      3768 GATGATGATGGAAGAGTGC 3785

RESULT 6
AA098221
ID   AA098221 standard; cDNA to mRNA; 7310 BP.
XX
AC   AA098221;
XX
DT   25-MAR-2003 (revised)
DT   15-AUG-1996 (first entry)
XX
DB   Hepatitis C virus clone genome.
XX
XX   Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
XX   antibodies; immunoprophylaxis; sera; serum; ds.
XX
OS   Hepatitis C virus.
XX
EN   US5443965-A.
XX
PD   22-AUG-1995.
XX
PF   05-APR-1991; 91US-00681703.
XX
PR   06-APR-1990; 90US-00505611.
PR   09-OCT-1990; 90US-00594854.
XX
PA   (GENE-) GENELABS INC.
XX
PI   Moeckli R, Reyes GR, Kim JP;
XX
DR   WPI; 1995-302120/39.
XX
PT   New nucleic acids encoding hepatitis C virus antigens - used to develop
PT   prods. for detection of HCV-infected sera and prodn. of vaccines and anti
XX   -HCV antibodies.
XX

```


PS Example 4; Fig 11; 71pp; English.

XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
CC sera and individuals infected with HCV. They can also be used in an anti-
CC HCV vaccine or for the production of anti-HCV antibodies which can be
CC used for passive immunoprophylaxis. The antigens consistently identify
CC more HCV positive serum samples with a high degree of specificity. See
CC AA098202-14 and AA81939-51. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PR field.)

SO Sequence 7310 BF; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,42e-241	Length:	7310
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AA098221 (1-7310)

```

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB ::::
DB 1728 CTGGCGCCCATCAAGCGGTACGCCAGCAAGGCGCTCTTAAGGTGATATAC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 1788 AGCTTAACCTGGCGGCAAAAACCAAGTGAAGGTAGGTCCAGATTGTGTCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 1848 GCCCAACCTTCTCGGCAACGTGCATCATATGGGTGTGTGCTGACGTACTACAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1908 GGAACGAGGACCACTGGGTACCCAGAGGTCCTGTCTCCAGATGTATCCAAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyValAlaArgSerLeuThrProCysThrCys 100
DB 1968 CAAAGCTTTGGGCTGGCCGCTCCGCAAGGTAGCGGTCTCATTTGACACCTGCACTTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 2028 GGCTCTCGGACCTTTACCTGTGTACAGAGGACGCCCATGTCTCCCGTGGCGCGGCGG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 2088 GGTGATGACAGGGGCGACCTGTGTCCGCCGCCCATTTCTTCACTTGAAAGGCTCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 2148 GGGGGTCCGCTGTGTGCCCGGGGCGACGCGTGGGCAATTTAAGGCGCGCGGTGCG 2207
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 2208 ACCCGTGAAGTGGCTTAAGCGGTGACCTTTATCCCTGTGAGAACTTAAGACACATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 2268 AGGTCCCCGGGTGTACAGGTAACCTCTCCACAGTAGTGCCCGAGAGCTTCCAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 2328 GCTCACTCCATGTCTCCACAGGCGGCAAAAGACCAAGGTCCCGGTGCAATAGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 2388 GCTAAGGGCTATTAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 2448 GCTTACATGTCTCAAGGCTCATGGATGATCTTAACATCAGAGCCGGGTGAGAAATTT 2507

```

```

QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 2508 ACCACTGGGAGCGCCCATCAAGTACTCCACCTACCGGCAAGTCTTCCGCGCGGGGTGC 2567
QY 281 SerGlyGlyAlaLysTrpAspIleIleCysAsnGluCysHisSerThrAspAlaThrSer 300
DB 2568 TCGGGGGCGCTTATACATTAATTAATTTGGAGAGTGTCCACTCCAGGTGCAATCC 2627
QY 301 IleLeuGlyIleGlyIleThrValLeuAspGlnAlaGlyIleThrAlaGlyLeuThrVal 320
DB 2628 ATCTTGGCATTCGGCACTGTCTTGAACAGAGACATCCGGGGCGAGACTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnIleVal 340
DB 2688 CTCGCCACCGCCACCCCTCGGGCTCGTCACTGTGCCCATCCCAATCGAGAGGTT 2747
QY 341 AlaLeuSerThrThrGlyGlyIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
DB 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATATC 2807
QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 2808 AAGGGGGGAGACATCTCATCTTCTGTATTCAAGAGAGTGCACAGAACTCGCCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 2868 AAGCTGGTGCATTGGGCAATCAATGCGGTGCTTACCGCGGTCTTGAAGTGTCCGCT 2927
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 2928 ATCCCAACGACCGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAlaAspSer 440
DB 2988 GGGCACTTGACTCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3047
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 3048 CTTGACCTTACCTTCAACATTTAGACATCAGCTCCCGGAGATGCTGTCCCGGACT 3107
QY 461 GlnArgArgGlyValArgThrGlyValArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 3108 CAAAGTGGGGGAGAGCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3167
QY 481 GlnArgProSerGlyLysPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 3168 GAGCGGCGCTCGGCATGTGTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3227
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 3228 GCTTGTGTAGCTCAAGCCCGCCGAGACATCAAGTTAAGGTAGAGGTACATGAACAC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 3288 CCGGGGCTTCCGCTGTGCGAGACATCTTGAATTTTGGAGGGCGCTTTTACAGGCGCTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrGlyGlnSerGlyGlnLeuProTyr 560
DB 3348 ACTCATATATATCCCACTTCTATCCCAACAAGAGAGTGGGAGAACTTCTTAC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB 3408 CTGGTAGCTTCAAGCCACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3467
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 3468 CAGATGTGAAGTGTGATTCCTCTCAAGCCCACTTCATGGGCAACACCTCTGTCTA 3527
QY 601 TyrArgGluGlyValAlaGlnAsnGlyValThrLeuThrHisProValThrLysTyrIle 620
DB 3528 TACAGACTGGGCGCTGTTCAGAAATGAATCACTTACAGCACCCAGTCAACCAATATCATC 3587

```

```
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValIleuValGlyGly 640
DB 3588 ATGCATGCGATGTCGGCGAGCTGAGCTCTCACAGCACTGGGCTGCTGTTGGCGGC 3647
QY 641 ValIleuAlaAlaIleuAlaAlaIleuValCysLeuSerThrGlyCysValIleValGlyArg 660
DB 3648 GTCTGCTGCTGCTTGGCGCGGATTTGCTGCATCAAGAGCTGCGTGCATATGTTGGGAGG 3707
QY 661 IleValIleuSerGlyLysProAlaIleIleProAspArgGluValIleuValArgGluPhe 680
DB 3708 GTCTGCTTGTCCGGGAAAGCCGCAATCATCTGACAGGAAAGTCTTACCGAGAGTTC 3767
QY 681 AspGluMetGluGluCys 686
DB 3768 GATGATGATGAGAGTGC 3785

RESULT 7
AAA75296
ID AAA75296 standard; cDNA; 8316 BP.
XX
AC AAA75296;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA sequence compiled Hepatitis C virus cDNA clones.
XX
KM Hepatitis C virus; HCV, antisense polynucleotide; polypeptide;
XX viral infectivity; viral replication; de.
XX
OS Hepatitis C virus.
XX
FH Key 1. 8316
FT CDS /*tag=
FT /note= "partial sequence; no termination codon given"
XX
PN EPI034785-A2.
XX
PD 13-SEP-2000.
XX
PF 16-MAR-1990; 2000EP-00109602.
XX
PR 17-MAR-1989; 89US-00325338.
XX PR 20-APR-1989; 89US-00341334.
XX PR 18-MAY-1989; 89US-00355002.
XX PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2000-566891/53.
XX DR P-PSDB; AAB18540.
XX
PT Novel composition comprising a hepatitis C virus antisense polynucleotide
XX PT which is complementary to or corresponds to a sense strand of the virus
XX PT genome, and selectively hybridizes to it.
XX
PS Example; Fig 16; 75pp; English.
XX
XX The specification describes a pharmaceutical composition which comprises
XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX characterized by a positive stranded RNA genome which has 40% homology at
XX the polypeptide level to a HCV polypeptide. The antisense polynucleotide
XX binds to cellular polynucleotides which enhance and/or are required for
XX viral infectivity, replicative ability or chronicity. The antisense
XX polynucleotides may also be designed to bind with high specificity, to be
XX of increased stability, to be stable and to have low toxicity. The
XX composition also comprises an agent which causes viral RNA to be
XX inactive. The composition is used for preventing HCV replication in a
XX system. The present sequence represents a novel HCV cDNA sequence, which
XX is used in the course of the invention
```

```
SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1 66e-241 Length: 8316
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75296 (1-8316)
QY 1 MetAlaProIleThrAlaTrpAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
DB 2734 CTGGCGCCATTCACGGGGTACGGCCAGACAGAAAGGGGCTCTTAAGGTGATTAATCACC 2793
QY 21 SerLeuThrGlyArgAspLysAlaGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 2794 AGCCTTAATGCGCGGACAAAACCAAGTGAAGGGTGAAGTCCAGATTGTCACTGCT 2853
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 2854 GCCCAAACTTCTCTGGCAAGTGCATCATGGGTGTCTGACCTGTACCAAGGGGCC 2913
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 2914 GGAACAGAGACATCGCGGTACCAAGGGCTGTGCATGCATGATATCAATATGATGAC 2973
QY 81 GluAspLeuValGlyTrpProAlaProGlnGlyValArgSerLeuThrProCysThrCys 100
DB 2974 CAAGACCTGTGTGGGTGGCCGCTCGCAAGTGAAGCGCTCATATGACACCTGCACCTGC 3033
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3034 GGCTCTCGAATCTTACCTTACCTGTCACAGAGCAAGCCGATCTATCCCGCGCCGGCG 3093
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 140
DB 3094 GTGATATGACAGGGGAGCGCTGTGTGCGCCCGCCATTTCTTAATGAAGGCTCTCG 3153
QY 141 GlyGlyProLeuLeuLeuProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3154 GGGGGTCCGCTGTGTGCCCCGCGGACAGCGGTGGCATATTAAGGCGCGGGTGTGC 3213
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGlnThrMet 180
DB 3214 ACCGTGAGTGGCTAAGGCGGTGACTTATCTCTGTGAGAACTGAGAACACATG 3273
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3274 AGGTCCCGGTGTTCACGAGTAACTCTCTCCACAGTAGTGCCTCCAGACTTCCAGGTG 3333
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3334 GCTCACTTCATCTCCACAGGACGAGGAGGAAAGCAACAGGTCCCGGTGATATGCA 3393
QY 221 AlaGlnGlyTyrLysValLeuValLeuAspProSerValAlaIleThrMetGlyPheGly 240
DB 3394 GCTCAAGGCTATTAAGGTGCTAGTACTCAACCTCTGTCTGTGCAACACTGGGCTTGGT 3453
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 3454 GCTTACATGTGCCAAGGCTCATGGATGATCTTAATCATCGACCGGGGTGAACAACATT 3513
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 3514 ACCACTGGAGCCCAACAGTACTCCACTACGCGCAAGTCTTGGCGAGGGGGTGC 3573
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 3574 TCGGGGGGCGCTTATGACATATATTTTGAAGAAGGCCACTCCACAGATGCAACATCC 3633
QY 301 IleuGluGlyIleGlyThrValIleuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
```

```

Db      3634 ATCTGGGCAATCGGCATGCTTGGACCAAGCAGACAGCTGGGGGGAACCTGGTTGTG 3693
Qy      321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      3694 CTGGCCACCGCCACCTTCCGGGCTCCGTCACTGTGCCCATCTCCAACTACAGAGAGATT 3753
Qy      341 AlaLeuSerThrThrGlyIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      3754 GCTCTGTCCACACCGAGAGATCTTTTATCGGCAAGGCTATCCCTCCGAAAGTATC 3813
Qy      361 LysGlyIleYargHisLeuIlePheCysHisSerLysLysCysAspGluLeuAla 380
Db      3814 AAGGGGGGAGACATCTCATCTTCTCATCAAGAAAGAGTGGCAGAACTGGCCGCA 3873
Qy      381 LysLeuValAlaLeuGlyValAspAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      3874 AAGCTGGTCGATGGGCATCAATGCGGTGGCTTACCGCGGTCTTGACGTGCTCCGTC 3993
Qy      401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      3934 ATCCGACCAAGCGCGCATGTTGTGTGGCAACGATGCCCTCATGACCGGCTATACC 3993
Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      3994 GGCACACTTCACCTCGGTGATGACTGCAATACGTGTGTCAACCCAGACAGTCAATTGACG 4053
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4054 CTGACCCCTACCTTACCATTTAGACAAATACGCTCCGCCAGAGATGCTGTCCGGCACT 4113
Qy      461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db      4114 CAACGTGGGGGCAAGGAGCTGGAGGGGGAAGCAGCAAGCATCTACATTTGGTGACCGGGG 4173
Qy      481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db      4174 GAGGCGCCCTCCGCGCATGTTGCACTCCGCTCTGTGATGCTATGACGACGAGCTGT 4233
Qy      501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuAspAlaTyrMetAspThr 520
Db      4234 GCTTGGATGAGCTACAGCCCGCCGAGACATCACTAGGCTACAGAGGTACATGAACCC 4293
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db      4294 CCGGGGCTTCCCGTGTCCAGAACCATCTTGAAATTTGGAGGGCGCTTTACAGGGCTC 4353
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysLeuProTyr 560
Db      4354 ACTCATATAGATGCCCACTTCTTATCCACAGCAAGAGAGTGGGAGAACTTCTTAC 4413
Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgLysAlaGlnAlaProProPheThrAsp 580
Db      4414 CTGGTACCGTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCTCCCACTCTGGGAC 4473
Qy      581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      4474 CAGATGTGGAAGTGTGATTCGCTCAAGCCCACTTCATGGGCCAAACCCCTCTTA 4533
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      4534 TACAGACTGGGCGCTGTTCAAGATGAATACCTCTGACGACCCAGTCAACAATATCATC 4593
Qy      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      4594 ATGCATGACATGTGCGCGGACCTGGAAGTGTCTCCAGGACCTGGGTGCTGTGGGGGC 4653
Qy      641 ValLeuAlaIleAlaValAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      4654 GTCTGTGCTGCTTGGCGCGGTATTTGCTGTCAACAGGCTCGGTGTCATATGTGGGAGG 4713
Qy      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680

```

```

Db      4714 GTCTCTTGTCCGGGAAGCCCGCAATCATCTGACAGGAGAGTCTTACGAGAGTTC 4773
Qy      681 AspGluMetGluGluCys 686
Db      4774 GATGAGATGGAAAGATGTC 4791

RESULT 8
AAZ07656
ID  AAZ07656 standard; DNA; 9133 BP.
XX
AC  AAZ07656;
XX
DT  20-MAR-2003 (revised)
DT  08-NOV-1999 (first entry)
XX
DE  Nucleotide sequence of HCV-1 ORF.
XX
KM  Hepatitis C virus; HCV; J1; J7; HCV-1; non-A; non-B HCV; NANBH;
KM  HCV infection; vaccine; ds.
XX
OS  Hepatitis C virus.
XX
FH  Key
FH  CDS
FT  Location/Qualifiers
FT  268..9132
FT  /*tag= b
FT  /transl_except= (pos:1588..1589; aa:Leu)
FT  /note= "this codon has an apparent 1 nucleotide deletion,
FT  which alters the reading frame"
FT  /transl_except= (pos:1647..1650; aa:Pro)
FT  /note= "this codon has an apparent 1 nucleotide
FT  insertion, which alters the reading frame; this insertion
FT  is not indicated in the sequence present in the formal
FT  sequence listing of the specification"
XX
PN  EP939128-A2.
XX
PD  01-SEP-1999.
XX
PF  17-SEP-1990; 99EP-00101746.
XX
PR  15-SEP-1989; 89US-00408045.
PR  21-DEC-1989; 89US-00456142.
PR  17-SEP-1990; 90EP-00310149.
XX
PA  (OYAA/) OYA A.
PA  (CHIR ) CHIRON CORP.
XX
PI  Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;
PI  Cha T, Irvine BD;
XX
DR  WPI: 1999-480843/41.
DR  P-PSDB; AAY14975.
XX
PT  New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PT  infections and development of vaccines.
XX
PS  Disclosure; Fig 12; 132pp; English.
XX
CC  The invention provides two new isolates of hepatitis C virus (HCV), J1
CC  and J7. These two isolates comprise nucleotide and amino acid sequences
CC  that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC  may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC  hybridisation for diagnosis of NANBH infections. They may also be used to
CC  screen blood donors, donated blood and blood products for this infection.
CC  The isolates may also be used to isolate other naturally occurring
CC  variants of the virus. The polypeptides may be used as a vaccine for
CC  administration to patients to protect against infection with NANBH. The
CC  present sequence represents the nucleotide sequence of HCV-1 ORF.
CC  (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC  correct PR field.)
XX
SQ  Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	1,866-241	Length:	9133
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
	2	Gaps:	0

US-09-930-591-2 (1-686) x AAZ07656 (1-9133)

```

QY      1 MetLaprollethrrAlaTyraAGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3343 CTGGGCGCCATCAGCGGCTACGCCACAGCAAGGGGCTCTTAAAGGTGCATATACCC 3402
QY      21 SerLeuThrgIyArqAspLyAsnGlnValGlnGlyValGlnIleValSerThra 40
Db      3403 AGCCTAATGCGCGGACAAAACCAAGTGGAGGTGAGTCCAGATTGTCAACGCT 3462
QY      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIYrHieGlyAla 60
Db      3463 GCCCAACCTTCTCGGCAACGTGCATCATGGGGTGGCTGAGCTGTCTACACGCGGCG 3522
QY      61 GlyThrArgThrIleAlaSerProIySerGlyProValIleGlnMetYrThrAsnValAsp 80
Db      3523 GGAAGAGAGACCATGCGCTCACCAAGGGTCTGTGCATCCAGATGTATACCAATGTAGAC 3582
QY      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3583 CAAGACCTTGGGGTGGCGCGCTCCGACAGGTAGCCGCTCATTGACACCTGTGACTTGC 3642
QY      101 GlySerSerAspLeuTyLeuValThrArgHieAlaAspValIleProValArgArgArg 120
Db      3643 GGCTCTCGAGACTTATACCTGTGTACGAGGACGCCCATGTCAATCCCGTCCGCGCGG 3702
QY      121 GLyAspGlyArqGlySerLeuLeuSerProArgProIleSerYrLeuLeuGlySerSer 140
Db      3703 GGTGATAGAGGGGAGCGAGCTGTGTGCGCCGCGCCATTTCTTAAGAGGCTCTCTCG 3762
QY      141 GlyGlyProLeuLeuLeuCysProAlaGlyHieAlaValGlyIlePheArgAlaAlaValCys 160
Db      3763 GGGGGTCCGCTGTGTGTGCGCGCGGAGCGCGCGGACATATTTAGGGCGCGGCTGTC 3822
QY      161 ThrArgGlyValAlaIleValAlaValAspPheIleProValGlnSerLeuGlnThrThre 180
Db      3823 ACCCTGTAGAGTGGCTTAAGCGGTGACTTATCCCTGTGAGAACTTAAGAGCAACATG 3882
QY      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
Db      3883 AGGTCCCGGTGTCAACGATTAACCTCTCTCCACAGTAGTGCCCAAGGCTTCCAGGTG 3942
QY      201 AlaHieLeuHieAlaProThrGlySerGlyLySerThrLySerValProAlaAlaTyra 220
Db      3943 GGTCACTTCATGCTCCACAGGACGCGCAAGGACCAAGGTCCCGCTGCATATGCA 4002
QY      221 AlaGlnGlyTyLeuValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240
Db      4003 GCTCAGGGGTATTAAGGTGTAGTACTCAACCCCTGTGTGCAAACTGGGGCTTTGGT 4062
QY      241 AlaTyMetSerSerLyAlaHieGlyIleAspProAsnIleArgThrGlyValAlaArgHie 260
Db      4063 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATAGACGCGGGTGAACAATT 4122
QY      261 ThrThrGlySerProIleThrTySerThrTyArgLyLePheLeuAlaAspGlyGlyCys 280
Db      4123 ACCACTGGGAGCCCACTACGTACTCACTCACTAGCGAACTTCTTCCGACGCGGGTGC 4182
QY      281 SerGlyGlyAlaTyAspIleIleIleCysAspGlyCysHieSerThrAspAlaThrSer 300
Db      4183 TCGGGGGGGGCTTATAGCATTAATATTGTGACAGATGCCACTCCACGATCCCAATCC 4242
QY      301 IleLeuGlyYrIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db      4243 ATCTTGGGATCGGCACTGTCTTGAACAAGACAGACTGCGGGGCGAGACTGGTTGTG 4302

```

```

QY      321 LeuAlaThrAlaThrProProGlySerValThrValProHieProAsnIleGlnGlyVal 340
Db      4303 CTGGCACCGGACCCCTCCGGGCTCGCTACTGTGGCCCATCCCAATCAGAGAGGTT 4362
QY      341 AlaLeuSerThrThrGlyGlnIleProPheTyGlyLyAlaIleProLeuGlnAlaIle 360
Db      4363 GCTGTGTCAACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAATAATATC 4422
QY      361 LyGlyGlyArgHieLeuIlePheCysHieSerLyAspLyAspGlyLeuAlaAla 380
Db      4423 AAGGGGGGAGACATCTCTCTGTGTCAATTAAGAAAGAGTGGAGAACTCCGCCCA 4482
QY      381 LyLeuValAlaLeuGlyValAlaAlaValAlaTyraGlyLyLeuAspValSerVal 400
Db      4483 AAGCTGTGTGCACTTGGGATCATATGCGGCTACTACCGCGGTCTTGACGTGCGCTC 4542
QY      401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      4543 ATCCGACACGCGGCGATGTGTGTGTGCGGACCGATGCCCTCATGACCGGCTATACC 4602
QY      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      4603 GCGACTTGACTCGATGTATACATGCAATAGTGTGTACCCAGAGCTGATTTACGC 4662
QY      441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4663 CTTGACCTTAACCTTCAACATTTAGACATACGCTCCCAAGATGCTGTCTCCGCACT 4722
QY      461 GlnArgArgGlyYrArgThrGlyArgGlyLyAspProGlyIleTyraArgPheValAlaProGly 480
Db      4723 CAACGTGCGGGGAGACTGCGAGAGGGGAAAGCAAGGATCTACAGATTTGTGGACCGGG 4782
QY      481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyCysTyraAspAlaGlyCys 500
Db      4783 GAGGCGCTCCCGGCAATGTTCAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4842
QY      501 AlaTrpTyrgIleLeuThrProAlaGlnThrThrValArgLeuArgAlaTyreAsnThr 520
Db      4843 GCTTGGATAGACTCAACGCGCGGAGACTACAGTTAGGCTACAGACGTACATGAACACC 4902
QY      521 ProGlyLeuProValCysGlnAspHieLeuGlnPheTrpGlnGlyValPheThrGlyLeu 540
Db      4903 CCGGGGCTTCCGTTGGCCAGACCACTTTGAATTTTGGAGGGCTCTTTACAGGCTTC 4962
QY      541 ThrHieLeuAspAlaHiePheLeuSerGlnThrLyGlnSerGlyGlnAsnLeuProTy 560
Db      4963 ACTCATATATATGCCCATCTTATCCCAAGCAAGAGTGGGAGAACTTCTCTTAC 5022
QY      561 LeuValAlaTyrgIlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db      5023 CTGGTACCGTAACAAGCACCGGTGTGCTAGAGGCTCAAGCCCTCCCATCTGTGGAC 5082
QY      581 GlnMetTrpLyCysLeuIleArgLeuLyProThrIleuHieGlyProThrProLeuLeu 600
Db      5083 CAGATGTGAAGATTTGATTCCTCTCAACCCACCTCATGGGCAACACCCCTGTGCTA 5142
QY      601 TyraArgLeuGlyAlaValGlnAsnGlnValThrLeuThrHieProValThrLyTyrIle 620
Db      5143 TAGAGACTGGGCGGTGTTCAGAAATGAATACCTTGACGACCAACCAATATCATC 5202
QY      621 MetThrCysMetSerAlaAspLeuGlnValValThrSerThrTrpValLeuValGlyGly 640
Db      5203 ATGACATGACATGTGGCGGACCTGAGAGTGTACGAGCACTGGGTGTCTGTGGCGGC 5262
QY      641 ValLeuAlaAlaLeuAlaAlaTyraCysLeuSerThrGlyCysValIleValGlyArg 660
Db      5263 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5322
QY      661 IleValLeuSerGlyLyAspProAlaIleIleProAspArgGlnValLeuTyraArgGlnPhe 680
Db      5323 GTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5382

```



```

Db      4655 GGGGACTTGCACTCGGTGATGAGACTGCATACGTTGTCACCCAGACAGTCATTTCCAGC 4714
Qy      441  LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4715 CTTGACCTTACCTTACCATTTAGAGACATACACGCTCCCGCAGATGCTGCTCCCGCACT 4774
Qy      461  GlnArgArgGlyArgThrGlyArgGlyValProGlyValLeuThrArgPheValAlaProGly 480
Db      4775 CAACGTGGGGGCGAGAGCTGGCGAGGGGAGACCGACATTCAGATTTGGTGGACCGGG 4834
Qy      481  GlnArgProSerGlyMetPheAspSerValLeuGlyGlyCysTyrAspAlaGlyCys 500
Db      4835 GAGGGCCCTCCGGCAGCTTGACTGCTCGTCTCTGTGAGTGTCTATGACCGCAGGCTGT 4894
Qy      501  AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      4895 GCTTGATGATGAGCTCACGCCCGCCGAGACATACAGTTAGGCTACAGAGGTTCATGAAACCC 4954
Qy      521  ProGlyLeuProValCysGlnAspHisLeuGluThrPheTrpGluGlyValPheThrGlyLeu 540
Db      4955 CCGGGGCTCCCGGTGCGAGACCACTTGAATTTGGAGGGCGCTTTTACAGGGCTTC 5014
Qy      541  ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db      5015 ACTCATATGATGCGCCACTTTCTATCCAGCAAGACAGAGTGGGAGAACTTCTTAC 5074
Qy      561  LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db      5075 CTGGTAGGTGACCAAGCACCGTGTGGCTGAGGCTCAAGGCTCCCTCCCATCGTGGGAC 5134
Qy      581  GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      5135 CAGATGTGAAGCTGTTGATTCGCTCAAGCCCAACCTCCATGGGCCCAACCCCTGCTA 5194
Qy      601  TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      5195 TACAGACTGGCGCGCTTCAAGATGAAATCACCCCTGAGCCACCAAGTCACCAATATCATC 5254
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyVal 640
Db      5255 ATGACATCATGTCGGCGACCTGAGGTCTGTACAGACACTGGGTCTGCTGGCGCG 5314
Qy      641  ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5315 GTCCGTGGTGGTGGCGCGGATTCGCTCAACAGGCTCGGTGATGATGATGATGATGATG 5374
Qy      661  IleValLeuSerGlyValProAlaAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
Db      5375 GTCCGTCTGTCTCGGGAGACCGGCATATCACTGACAGGGAAGTCTCTACCGAGAGTTC 5434
Qy      681  AspGluMetGluGluCys 686
Db      5435 GATGAGATGAGAGAGTGC 5452

RESULT 10
AAA75297
ID      AAA75297 standard; cDNA; 9185 BP.
XX
AC      AAA75297;
XX
DT      15-JAN-2001 (first entry)
XX
DE      Sense strand of HCV encoding a polypeptide.
XX
KM      Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;
XX      viral infectivity; viral replication; ds.
OS      Hepatitis C virus.
XX
Key      Location/Qualifiers
FH      CDS
FT      320..9184
FT      /tag=a
FT      /note="partial sequence; no termination codon given"

```

```

XX      XX
PN      EP1034785-A2.
XX      13-SEP-2000.
XX      16-MAR-1990; 2000EP-00109602.
XX      17-MAR-1989; 89US-00325338.
XX      20-APR-1989; 89US-00341334.
XX      18-MAY-1989; 89US-00355002.
XX      16-MAR-1990; 90EP-00302866.
XX      (CHIR ) CHIRON CORP.
XX      Houghton M, Choo Q, Kuo G;
XX      WPI, 2000-566891/53.
XX      P-PSDB; AAB18541.
XX      Novel composition comprising a hepatitis C virus antisense polynucleotide
XX      PT      which is complementary to or corresponds to a sense strand of the virus
XX      PT      genome, and selectively hybridizes to it.
XX      Example, Fig 17; 75pp; English.
XX      The specification describes a pharmaceutical composition which comprises
XX      CC      a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX      CC      characterized by a positive stranded RNA genome which has 40% homology at
XX      CC      the polypeptide level to a HCV polypeptide. The antisense polynucleotide
XX      CC      binds to cellular polynucleotides which enhance and/or are required for
XX      CC      viral infectivity, replicative ability or chronicity. The antisense
XX      CC      polynucleotides may also be designed to bind with high specificity, to be
XX      CC      of increased stability, to be stable and to have low toxicity. The
XX      CC      composition also comprises an agent which causes viral RNA to be
XX      CC      inactive. The composition is used for preventing HCV replication in a
XX      CC      system. The present sequence represents a novel HCV cDNA sequence, which
XX      CC      is used in the course of the invention
XX      SQ      Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.88e-241 Length: 9185
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x AAA75297 (1-9185)
Qy      1  MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3395  :::: 3574.00
Qy      21  SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValIleSerThrAla 40
Db      3455  AGCTTAATCGCCCGGACCAAAACCAAGTGAAGGTAGATCCAGATGTGTCAACTGCT 3514
Qy      41  AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      3515  GCCCAAACTTCTCGGCACAGTCGATCAATGGGGTGTGTGACTGTCTACACAGGGGCC 3574
Qy      61  GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrTrpAsnValAsp 80
Db      3575  GGAACGAGGACCAATCGGTGTCACCCAGAGGTCCTGTCTATCCAGATGTATACCATGTGAC 3634
Qy      81  GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysTrpCys 100
Db      3635  CAAGACTTGTGGCTGGCCGCTCCGCAAGTGAAGCGGCTCATTTGACACCCCTGCATCTGC 3694
Qy      101  GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      3695  GGCTCTCGGACCTTTACTGTGTGACAGGAGCAGCCGATGTCAATTCCTGTCGCCGCGCG 3754

```

OY	121	GIYAspGlyAlaArgGlySerLeuIleuSerProAlaProIleSerTyrLeuIleuGlySerSer	140
Db	3755	GATGATACAGAGGGGAGCGCTGTGTCCGCCCGGCCCATTTCTACTTGAAGAGCTTCCTCG	3814
OY	141	GIYGIYProLeuIleuCyAsProAlaGlyHisAlaValaGlyIlePheArgAlaAlaValaCys	160
Db	3815	GGGGGTCCGGCTTTGTGCCCGCGGGGACCGCGTGGGCAATTTTAGGGCCCGGGTGTGC	3877
OY	161	ThrArgIlyValAlaLeuAlaValaIAspPheIleProValGluSerLeuGluThrThyMet	180
Db	3875	ACCCCTGAGATGGCTTAAGCGCGGTGACATTATCTCTGGAGAACTTAGAGACAACACTAG	3934
OY	181	ArgSerProValPheSerAspAsnSerSerProAlaValProGluSerTyrGlnIval	200
Db	3935	AGGTCCCGAGTGTTCAGGATTAACCTCTCCACCAAGTAAGTCCCAAGCTTCAGAGTGG	3994
OY	201	AlaHisLeuHisAlaProThrGlySerGlyIysSerThrIlyValProAlaAlaTyrAla	220
Db	3995	GCTCACCTCCATGCTCCACAGGACGGGCAAAAGACCAAGAGTCCGGCTGCATATGCA	4055
OY	221	AlaGlnGlyTyrIlyValIleuValIleuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4055	GCTCAGGGCTAATAAGTGTAGTACTCAACCCCTGTGTTGCTGAACAACCTGGCTTGTGT	4114
OY	241	AlaTyrMetSerIlyValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	4115	GCTTACATAGTCCMAAGGCTCATGGATGCACTCTTAACATCAGGACCGGGGTGAGAACAT	4174
OY	261	ThrThrIysSerProIleThrTyrSerThrTyrGlyIysPheLeuAlaAspGlyGlyIys	280
Db	4175	ACCATGGCAGGCCCATCGATCGTACTCCACCTACGGCAAGTCTTGTCCGACGGCGGGTGC	4233
OY	281	SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4235	TGGGGGGGGCGTTAAGACATAATATTGTGACGAGTCCACTCCACGGATCCCACTCC	4294
OY	301	IleIleuGlyIleGlyThrValIleuAspGlnaGluThrAlaGlyValaArgLeuThrVal	320
Db	4295	ATCTTGGGCATCGGACATGCTCTTGACCAAGAGAGACTGGCGGGGACAGACTGGTGTG	4354
OY	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
Db	4355	CTGGCACCGGCACCCCTCCGGGCTCCGTCATGTGGCCCAATCCCAATCCAGAGAGT	4414
OY	341	AlaIleuSerThrThrGlyGluIleProPheTyrGlyIysAlaIleProLeuGlnIvalIle	360
Db	4415	GCTGTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATTC	4474
OY	361	LysGlyGlyIaArgHisLeuIlePheCysHisSerIlyIysIlyCysAspGluIleuAlaAla	380
Db	4475	AAGGGGGGAGACATCTCATCTTCTGTCAATTCMAAGAAAGTGGAGCAACTGCCGCA	4534
OY	381	LysLeuValAlaIleuGlyValaAsnAlaValaAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4535	AAGCTGTGCAATTTGGGATCATATCCGTGGCTTACTACCGGGCTTGAACGTGCCCTC	4594
OY	401	IleProThrSerGlyAspValValValaValaAlaThrAspAlaIleMetThrGlyPheThr	420
Db	4555	ATCCCGACCAAGGAGATGTTGCTCGTGGCAACGATGCCCTCATAGACGGCATATACC	4655
OY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValaThrGlnThrValaAspPheSer	440
Db	4655	GGCGACTTCGACTCGGTATGACGTCAATACGTGTGTACCCCAACACAGTGCATTTCAGC	4714
OY	441	LeuAspProThrPheThrIleGluThrIleThrIleuProGlnAspAlaValSerArgThr	460
Db	4715	CTTGACCTTACCTTCACCATTTGAGCAATACAGCTCCCAAGAAAGCTGTCTCCGCACT	4774
OY	461	GlnaAspArgIlyArgThrGlyIaArgIlyIysProGlyIleTyrArgPheValAlaProGly	480
Db	4775	CAACCTCGGGGAGGACTGGCGGGGGGAAAGCCAGGCAATTCAGATTTGTGTGACCGGGGG	4834

QY	481	GLNLRPR	ProSerGlyMetPheAspSerValLeuCYsGluCYsTYRAspAlaGlyCYs	500
Db	4835	GAGGGCCCCCTCCGGCAATGGTGGACTGTCGCTCTGTGAAGTCATAGACGAGGCTGT		4894
QY	501	AlaTrpTYRGLueuThrProAlaGluThrThrValArgLeuArgAlaTYRMetAsnThr		520
Db	4895	GCTTGATGATAGCTCAGAGCCCGGAGACATACATTAGGGCTAGAGGCTACATGAACCC		4954
QY	521	ProGlyLeuProValCYsGlnAspHisLeuGluPheThrPGLuGlyValPheThrGlyLeu		540
Db	4955	CCGGGGGCTTCCCGTGGTCCAGAGACCATCTTGAAATTTGGAGGGCCGCTTTTACAGGGCTC		5014
QY	541	ThrHisIleAspAlaHisPheLeuSerGlnThrThrIleGlnSerGlyIleAsnLeuProTYR		560
Db	5015	ACTCATATAGATGGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAAACCTTCCCTTAC		5074
QY	561	LeuValAlaTYRGlinalaThrValCYsAlaArgAlaGlnAlaProProSerTrpAsp		580
Db	5075	CTGGTACCGTACCAAGCCACCGTGTGGGCTTAGGGCTGAACCCCTCCCATGCTGGGAC		5134
QY	581	GlnMetTrpIleCYsLeuIleIleArgLeuIleProThrLeuHisGlyProThrProLeuLeu		600
Db	5135	CAGATGTGAAGTGTATTGATTGGCTCAAGCCCACTCCATGGGCAACACCCCTGTCTA		5194
QY	601	TYRTrpGluGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrIleTYRILE		620
Db	5195	TACAGACTGGGGCGGTGTTCAGAAATGAATACCCCTGACGACCCAGTCAACCAATACATC		5254
QY	621	MetThrCYsMetSerIleAspLeuGluValValThrSerThrTrpValLeuValGlyGly		640
Db	5255	ATGACATGCAATGTGGCCGACCTGGAGGTGTCTACAGACCACTGGGGTCTGTTGGCGGC		5314
QY	641	ValLeuAlaAlaLeuAlaAlaAlaTYRCYsLeuSerThrGlyCYsValValIleValGlyArg		660
Db	5315	GTCCTGGCTGCTTTGGCCGGCTGATGGCTGTCAACAAGGCTGGGTGTCATATGAGGGCAGG		5374
QY	661	IleValLeuSerGlyIleProAlaIleIleProAspArgGluValLeuTYRArgGluPhe		680
Db	5375	GTCCTGTGTTCGGGGAAGCCGGCAATCATCTGACAGGGAAGTCTTACCGAGAGTTC		5434
QY	681	AspGluMetGluGluCYs 686		
Db	5435	GATGAGATTGGAGAGTGC 5452		
RESULT 11				
AAT12710				
ID	AAT12710	standard; cDNA; 9401 BP.		
AC				
XX	AAT12710;			
XX				
DT	25-MAR-2003	(revised)		
DT	15-MAY-1996	(first entry)		
XX				
DE	Hepatitis C virus polyprotein.			
XX				
KW	Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;			
KW	antibodies; ds.			
XX				
OS	Hepatitis C virus.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	342..9378		
FT	/*Cag= a			
XX	EP693687-A1.			
XX				
PD	24-JAN-1996.			
XX				
PF	03-APR-1991;	95BP-00114016.		
XX				
PR	04-APR-1990;	90US-00504352.		
XX				
PA	(CHIR) CHIRON CORP.			

XX Houghton M, Choo Q, Kuo G;
XX WPI, 1996-117956/13.
DR P-PSDB; AAR90931.
XX
PT Combinations of synthetic Hepatitis C Virus antigens - provide more
PT effective diagnosis of Non-A, Non-B Hepatitis.
XX
PS Disclosure; Fig 1(A-Y); 53pp; English.
XX
CC The combination comprises an HCV antigen from the C domain (pref. C22 -
CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
CC (AAR90934) domain. The antigens may in the form of a fusion protein, a
CC simple physical mixture, or the individual antigens commonly bound to a
CC solid matrix. They are pref. prepd. by recombinant DNA techniques
CC (primers are given in AAT12711-T12716), but can be synthesised or
CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
CC to correct PF field.)
XX

SO Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,93e-241	Length:	9401
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AAT12710 (1-9401)

Qy 1 MetLapProIleThraIaYrAlaGlnGlnThraRgIyLeuLeuGlyCyAlIeIeThr 20
Db 3417 CTGCGCCCATCAGCGCTACGCCAGACAGAGGCGCTCTTGAAGGTGATATATACC 3476
Qy 21 SerLeuThrGlyARgAspLyAsnGlnValGlnGlyGlnValGlnIleValSerThrAla 40
Db 3477 AGCTTAATGCGCGGAGCAAAACCAAGTGAAGGTGAGTCCAGATTGTCTCAACGTCT 3536
Qy 41 AlaGlnThrPheLeuAlaIaThrCyAlIeAsnGlyValCyStrpThrValIyrHiGlyAla 60
Db 3537 GCCCAAACTTCTCGGCAACGTGCATCATGCGGTGCTGAGCTGTCTTACACAGCGGCGC 3596
Qy 61 GlyThrArgTrIleAlaSerProLySgLyProValIleGlnMetYrThAsnValAsp 80
Db 3597 GGAACGAGACCATCGCTCACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTTPProIaIaProGlnIyAlaArgSerLeuThrProCyStrhCys 100
Db 3657 CAAGACTTGTGGCTGGCGCGCTCCGCAAGTGAAGCGCTCATTTAGACACCTGCACCTTGC 3716
Qy 101 GlySerSerAspLeuYrLeuValIaThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGCTTCCTCGACCTTTACCTGTGACGAGGACCGCATTCATTCCTGCGCGCGCGG 3776
Qy 121 GlnAspGlyARgGlySerLeuLeuSerProArgProIleSerYrLeuLyGlySerSer 140
Db 3777 GGTATAGCAGGGGAGCGCTGTGTGCGCGCGCGCATTTCTTCTTGAAGGCTCTCTCG 3836
Qy 141 GlnGlyProLeuLeuCySProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGCTCGCTGTGTGTGCGCGCGGAGCACCGCGGCGCATATTTAAGGGCGCGGTGTC 3896
Qy 161 ThrArgGlyValAlaIaValAlaValAspPheIleProValGlnSerLeuGlnThrThMet 180
Db 3897 ACCGTGAGTGGCTTAAGCGGTGACTTATCTCTGTGAGAACTTAAGAACACCATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
Db 3957 AGGTCCCGGTGTTCACGATTAATCTCTCTCCACCAAGTATGTCGCCAGAGCTTCCAGGTG 4016

Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLySerThrIyValProAlaAlaIyrAla 220
Db 4017 GCTCACCTCATCTGCCACAGGACGCGCAAAAGCACCAAGGTCCGGCTGCATATGCA 4076
Qy 221 AlaGlnGlyYrIyIyValLeuValIleuAnProSerValAlaIaIaThrMetGlyPheGly 240
Db 4077 GCTCAGGCTATAGAGGTGCTAGTACACCCCTGTGTGCTGCAACACTGGGCTTGTGT 4136
Qy 241 AlaYrMetSerIyAlaHisGlyIleAspProAsnIleArgThrGlyValaArgThrIle 260
Db 4137 GCTTACATGTCCAAAGCTCATGGGATCGATCCCTAACATCAGACCGGGGTGAACAAT 4196
Qy 261 ThrThrGlySerProIleThrYrYSerThrYrGlyLyPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGACGCCCATCAAGTCTCACCTACCGGAAGTTCCTTCCGACGCGGGGTGC 4256
Qy 281 SerGlyGlyValaIyYrAspIleIleIleCysAspGlnCyHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGGCGCTTATGACATTAATTTTGACGAGTGCACATCCAGATGCCACATTC 4316
Qy 301 IleLeuGlyIleGlyThrValIleuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCAATCGGCACTGTCTTGAACAGACAGAGACTCGGGGGAGACTGTTGTG 4376
Qy 321 LeuAlaThraIaThrProProGlySerValIaThrValProHisProAsnIleGlnGlnVal 340
Db 4377 CTGCGCACCGCCCACTCCCTCGGGCTCGCTCATCTGTGCGCCATCCCAACATCGAGAGGT 4436
Qy 341 AlaLeuSerThrThrArgIyIleProPheYrGlyLyAlaIleProLeuGlnAlaIle 360
Db 4437 GCTGTGTCCACCAACCGAGAGATCCCTTTTACGGAAGGCTATCCCTCCGAAAGTATTC 4496
Qy 361 LySgLyGlyYrArgHisLeuIlePheCyHisSerIyLySgLyCysAspGlnLeuAlaIa 380
Db 4497 AAGGGGGGAGACATCATCTCTGTGCATTAAGAAAGAGTGGCGCAACTCGCGCA 4556
Qy 381 LySLeuValAlaLeuGlyValaAsnAlaValaIaIyYrYrArgGlyLeuAspValSerVal 400
Db 4557 AAGCTGTCCCATTTGGGCAATCAATGCGGTGCTTACTACCGCGTCTTGACGTGTCCGTC 4616
Qy 401 IleProThrSerGlyAspValValaIaValaIaThrAspAlaLeuMetThrGlyPheThr 420
Db 4617 ATCCGACACGCGCGCATGT 4676
Qy 421 GlyAspPheAspSerValIleAspCyAsnIleThrCysValIaThrGlnThrValaAspPheSer 440
Db 4677 GCGGACTTGCATCGGATGATGACCTGCATATACGATGTGTGTGTGTGTGTGTGTGTGT 4736
Qy 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4737 CTTGACCTTACCTTACCATTTGAGACAAATACGCTCCCGAGATGCTGTCTCCGCACT 4796
Qy 461 GlnArgArgGlyArgTrIleArgGlyLyPProGlyIleYrArgPheValaIaProGly 480
Db 4797 CAACGTGGGGGAGGACTGCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4856
Qy 481 GlnArgProSerGlyMetPheAspSerSerValIleuCySgLyCysYrAspAlaGlyCys 500
Db 4857 GAGGCGCCCTCGGCAATGTGCACTCGTCCGTCTGTGAGTGTATGACGAGGCTGT 4916
Qy 501 AlaTrpYrGlnLeuThrProAlaGlnThrThrValaArgLeuArgAlaIyrmMetAnThr 520
Db 4917 GCTTGATATAGCTCAAGCGCGGAGACTTACAGTTAGGTACGAGGTTACATGAACACC 4976
Qy 521 ProGlyLeuProValCySglnAspHisLeuGlnPheTrpGlnGlyValaPheThrGlyLeu 540
Db 4977 CCGGGGCTTCCGTGTGCGGAGACATCTTGAATTTTGGAGGGCGCTTTTACAGGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIySglnSerGlyGlnLeuProYr 560
Db 5037 ACTCATATATAGTCCCACTTTATCCACAGAAAGAGAGTGGGAGAACCTTCCCTTAC 5096
Qy 561 LeuValAlaIyYrGlnAlaIaThrValCySAlaArgAlaGlnAlaIaProProSerTrpAsp 580


```

Db      5097 CTGTAGCTTACCAAGACCGTGTGCTTAGGGCTCAAGCCCTCCCATCGTGGAC 5156
Qy      581  GIMeETrPhyScySeuileArSgLeuLysProThrluHsiGlyProThrProLeuL 600
Db      5157 CAGATGTGAAGGTTTGAATTCGCTTCAAGCCCACTCCATGGGCCAACACCCCTGCTA 5216
Qy      601  TTAAGTgLeuGlyAlaValGlnAsnGluValThrluThrluHsiProValThrLys 620
Db      5217 TACGACTGGGCGGCTGTTCAAGATGAATCATCCCTGACGACCCAGCACCAATAATCATC 5276
Qy      621  MetThrCysMetSerAlaAspLeuGluValValThrluSerThrluValLeuValGly 640
Db      5277 ATACATGATGATGTCGGCGGACCTGGAGAGTGTCTACGAGCACCTGGGCTGTTGGCGGC 5336
Qy      641  ValLeuAlaAlaLeuAlaAlaValCysLeuSerThrluGlyCysValValLeuValGly 660
Db      5337 GTCTGTGCTCTTGTGGCGGGATGTCCTGTCAACAGGCTGGCGTCAATAGTGGGAGG 5396
Qy      661  lLeValLeuSerGlyLysProAlaAlaLeuProAspArgGluValLeuTyrArgGluPhe 680
Db      5397 GTGTCTTGTCTCGGGAAGCGGCAATCATCTGACAGGAAGTCTTACGAGAGTTC 5456
Qy      681  AApGluMetGluGluCys 686
Db      5457 GATGATGATGGAAGTGC 5474

RESULT 12
AAT99981
ID      AAT99981 standard; DNA; 9401 BP.
AC      AAT99981;
XX
XX      25-MAR-2003 (revised)
DT      16-MAR-1998 (first entry)
XX
XX      HCV polyprotein coding sequence.
DE
XX      PCR primer; amplify: HCV; hepatitis c virus; antigen combination; NS3;
KM      C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW      NS4; de.
XX
XX      Hepatitis C virus.
OS
XX
XX      Key      Location/Qualifiers
FT      CDS      342..9377
FT      FT      /*tag= a
XX
XX      US5683864-A.
PN
XX      04-NOV-1997.
PD
XX      07-JUL-1992; 92US-00910760.
PE
XX
XX      18-NOV-1987; 87US-00122714.
PR      30-DEC-1987; 87US-00139886.
PR      26-FEB-1988; 88US-00161072.
PR      06-MAY-1988; 88US-00191263.
PR      26-OCT-1988; 88US-00263584.
PR      14-NOV-1989; 89US-00271450.
PR      17-MAR-1989; 89US-00325338.
PR      20-APR-1989; 89US-00341334.
PR      21-APR-1989; 89US-00353896.
PR      18-MAY-1989; 89US-00355002.
PR      04-APR-1990; 90US-00504552.
XX
XX      (CHIR ) CHIRON CORP.
PA
XX      Kuo G, Houghton M, Choo Q;
PI
XX      WPI, 1997-548976/50.
DR      P-PSDB; AAM34480.
XX

```

```

PT      Combination of three hepatitis C virus antigens - used for detection of
XX      specific antibodies to diagnose infection.
XX
XX      Disclosure; Col 25-46; 57pp; English.
XX
CC      This sequence represents the Hepatitis C virus polyprotein coding
CC      sequence. Fragments of this sequence can be amplified and used in the
CC      combination of HCV antigens of the invention. The HCV antigen combination
CC      comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC      1-120 of the HCV polyprotein), or its immunologically reactive fragment
CC      containing at least 8 aa. It also comprises two additional antigens from
CC      two different polyprotein domains, including at least 8 aa from the NS3,
CC      NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to
CC      aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.
CC      Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC      regions of the HCV polyprotein. These antigen combinations are used
CC      diagnostically to detect anti-HCV antibodies, using any standard
CC      immunoassay format. These antigen combinations have a broader range of
CC      reactivity with antibodies than any antigen individually. (Updated on 25-
CC      MAR-2003 to correct PR field.)
XX
SQ      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1,93e-241      Length:      9401
Score:      3574.00      Matches:      672
Percent Similarity:      99.428      Conservative:      10
Best Local Similarity:      97.968      Mismatches:      4
Query Match:      98.788      Indels:      0
DB:      2      Gaps:      0

US-09-930-591-2 (1-686) x AAT99981 (1-9401)
Qy      1      MetAlaProIleThrLarLarAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3417 CTGGCCCATCAACGCGGATGCGCCGACGACCAAGGGGCGCTCTAGAGTCAATAACACC 3476
Qy      21      SerLeuThrGlyYArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3477 AGCTTAACGTGGCGGAGCAAAACCAAGTGAAGGGTCCAGATTGTGTCAACTGCT 3536
Qy      41      AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleYrHsiGlyAla 60
Db      3537 GCCCAAACTTCTCTGCAACAGTGCATCAATGGGGTGTGCTGACGTCTAACACCGGGGCC 3596
Qy      61      GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetYrThrAsnValAsp 80
Db      3597 GGAACGAGGACATCGCGTCAACCAAGGCTCTGTATCCAGATGATACCAATGTAGAC 3656
Qy      81      GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3657 CAGACCTTGTGGGCTGGCGCCGCTCCGCAAGTAGCCGCTCATTTGACACCTGCACTTGC 3716
Qy      101      GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      3717 GGCTCTCGGACCTTAACTTACGTGTCAGAGGACGCGGATGTATCCCGGCGCGGGCGG 3776
Qy      121      GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLysGlySerSer 140
Db      3777 GGTGATAGCAGGGGCGCTGCTGCTGCCCGCCCAATTTCTCAATGTGAAGGCTCTCGG 3836
Qy      141      GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      3837 GGGGGTCCGCTGTGTGGCCCGGGGACGCGGTGGGATATTATTTGGCGCGGCTGTGC 3896
Qy      161      ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGlnThrThrMet 180
Db      3897 ACCGTGAGAGTGGCTAAGCGGCTTATTCCTGTGAGAGAACTTAAGACCAACCATG 3956
Qy      181      ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
Db      3957 AGGTCCCGGATTTCAAGATTAATCTCTCTCCACGAGTAGTGCCCAAGAGCTTCAGGTG 4016

```

QY 201 AlaHisLeuHiSaLaPProThrGlySerGlyIlySerThrIlyValProIlaAlaIyAla 220
 DB 4017 GCTCACCCTTCATGCTCCACAGAGCGGCAAAAGCAACCAAGGTCCTGGCTGATAGCA 4076
 QY 221 AlAGInGIyTyRlyValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 4077 GCTCAGGGGCTTAAGAGTGTAGTACTCAACCCCTCTGTCTGTGACACCTGGGCTTGGT 4136
 QY 241 AlaTyMeSerIlySerValaHISgLyIleAspProAsnIleArgThrGlyValaGlyThrIle 260
 DB 4137 GCTTACATGTCACAGGCTCATGGAGATCATCTTAACATCAGAGCCGGGGTGAGAACATTT 4196
 QY 261 ThrThrGlySerProIleThrIlySerThrIlyGlyIlyPheLeuAlaAspGlyGlyCys 280
 DB 4197 ACCACTGACAGCCCATCAGACTCCACCTACGCGCAAGTTCCTTGGCCGACGGGGGTGC 4256
 QY 281 SerGlyGlyAlaTyRAspIleIleIleCysAspGlyCysHISerThrAspAlaThrSer 300
 DB 4257 TCGGGGGGGCGCTTATGACATATTAATTGTGACGAGTGCACCTCCAGATGCCACATCC 4316
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValaArgLeuThrVal 320
 DB 4317 ATCTTGGGCGATCGGACACTGCTCTTGACCAAGACAGACTGGGGGGGAGACTGCTGTG 4376
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 4377 CTGCGCACCGGCAACCCCTCCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGAGATT 4436
 QY 341 AlaLeuSerThrThrGlyGlyIleProPheTyRlyIlyAlaIleProLeuGluAlaIle 360
 DB 4437 GCTCTGTCCACCAACGAGAGATCCCTTTTACGCGAAGGCTATCCCTCCGAAAGTATC 4496
 QY 361 LysGlyGlyIyArgHISLeuIlePheCysHISerIlyIlyIlyCysAspGlyLeuAlaAla 380
 DB 4497 AAGGGGGGAGACATCTATCTCTGTATTCAGAAAGAGTGACAGAACTGGCCGCA 4556
 QY 381 LysLeuValAlaLeuGlyValaAsnAlaValaIyTyRArgGlyLeuAspValSerVal 400
 DB 4557 AAGCTGTGCGATTCGGATCATAGTCGGTGCATACCCGCGCTTGAAGTGTCCGTC 4616
 QY 401 IleProThrSerGlyAspValValaValaValaIaThrAspAlaLeuMetThrGlyPheThr 420
 DB 4617 ATCCCGACGAGCGGAGTGTGTCGTGCGCAACGATGCCCTCATACCGGCTATACC 4676
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIThrGlnThrValaAspPheSer 440
 DB 4677 GGGGACTTCGACTCGGTATGACTGCATAGCTGTGTACCCAGACAGCTGATTTGAC 4736
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValaSerArgThr 460
 DB 4737 CTTGACCTTACCTTACCATTTAGACAATACGCTCCCCAGAGTGTGTCTCCGCACT 4796
 QY 461 GlnaGlyArgIyArgThrGlyIyArgGlyIySerProGlyIleTyRArgPheValAlaProGly 480
 DB 4797 CAACGTCGGGAGAGACTGAGCGGAGGAGCAAGCACTTCAGATTGTGTGACACCGAG 4856
 QY 481 GluArgProSerGlyMeCpHeAspSerSerValleuCysGlyCysTyRAspAlaGlyCys 500
 DB 4857 GAGCGCCCTCCGCGATGTGACCTGTCGCTCTGTGATGTGTATGACGACAGCTGT 4916
 QY 501 AlaTyRlyGluLeuThrProAlaGluThrThrValaGlyLeuArgAlaTyRMetAsnThr 520
 DB 4917 GCTTGGTATGAGCTACGCGCCCGAGACTACAGTTAGGCTTACGAGCGTACATGAACCC 4976
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValaPheThrGlyLeu 540
 DB 4977 CCGGGGGCTTCGGTGGCAGAGCACTTGAAATTTGGAGAGGCGTCTTTACAGGCTTC 5036
 QY 541 ThrHISileAspAlaHisPheLeuSerGlnThrIyGlnSerGlyIlyLeuAsnLeuProTyR 560
 DB 5037 ACTCANATAGATGCCCACTTTCTATCCAGAACAGAGTGGGAGAACTTCTTAC 5096
 QY 561 LeuValAlaTyRAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580

DB 5097 CTGTAGGCTACCAACCAACCGCTGTGGCTTGGGCTCAAGCCCTCCCATCGTGGAGC 5156
 QY 581 GlnMetTrpIyCysLeuIleArgLeuIlySPProThrLeuHISgLyProThrProLeuLeu 600
 DB 5157 CAGATGTGAATGTGTGATTGCTTCCAGGCCCACTTCATGAGGCAACACCTGTGCTA 5216
 QY 601 TyRArgLeuGlyAlaValGlnAsnGluValThrLeuThrHISProValThrIyTyRle 620
 DB 5217 TACAGACTGGGGCGCTTTCAGATGAATACCTTCAGACCACTGACCAATAATACATC 5276
 QY 621 MetThrCysMetSerAlaAspLeuGluValaIleThrSerThrTyRValLeuValGly 640
 DB 5277 ATGACATGATGTGGCCGACCTGAGGTGTCTACAGAGCACTGGAGTGTCTGTGGGCG 5336
 QY 641 ValLeuAlaAlaLeuAlaAlaTyRcysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 5337 GTCTTGCTGCTTGTGGCCCGCGATTCCTGTCAACAGCTGTGCTGTATGTGGGACAG 5396
 QY 661 IleValLeuSerGlyIyAspProAlaIleIleProAspArgGluValLeuTyRArgGlyPhe 680
 DB 5397 GTGCTTGTTCGGGAAAGCCGCAATCATACCTGACAGGGAAGTCTCTTACCGAGATT 5456
 QY 681 AspGluMetGluGluCys 686
 DB 5457 GATGAGATGGAAGAGTGC 5474
 DB
 RESULT 13
 AAV09989
 ID AAV09989 standard; DNA; 9401 BP.
 AC AAV09989;
 XX 26-MAY-1998 (first entry)
 DT XX
 DE HCV polyprotein coding sequence DNA.
 XX
 KW Hepatitis C virus C domain; HCV; C antigen; immunological activity;
 KM NS3 domain; NS4 domain; S domain; NS5 domain; de.
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT 342..9377
 FT CDS /*tag= a
 FT /product= "HCV polyprotein"
 XX
 PN US5712087-A.
 XX
 XX 27-JAN-1998.
 FD
 XX 12-MAY-1995; 95US-00440519.
 PF
 XX 04-APR-1990; 90US-00504352.
 PR 07-JUL-1992; 92US-00910760.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Kuo G, Houghton M, Choo Q;
 XX
 DR WPI: 1998-119973/11.
 DR P-PSDB: MAM40038.
 XX
 PT Immunoassays for hepatitis C virus antibodies - using combinations of
 PT antigenic fragments of HCV polyprotein.
 XX
 PS Disclosure; Fig 1; 59pp; English.
 XX
 CC This sequence encodes the hepatitis C virus (HCV) polyprotein which is
 CC used in the construction of novel combinations of HCV antigens that have
 CC a broader range of immunological activity than any single HCV antigen. An
 CC example of such an antigen given in this specification comprises a first
 CC antigen containing at least 8 amino acids of the C domain of the HCV

CC polypeptide and a second antigen comprising at least 8 amino acids of the
 CC NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV
 CC polypeptide in the form of a fusion protein, a physical mixture or bound
 CC to a solid matrix

XX Sequence 9401 BP, 1883 A, 2860 C, 2673 G, 1985 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	1,936-241	Length:	9401
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	2	Gaps:	0

US-09-930-591-2 (1-686) x AAV09989 (1-9401)

QY	1	MetaLapProilethralaYrAlaGlnGlnThraArglyLeuleuglyCysilelethr	20
DB	3417	CTGGCGCCCATCAAGCGGTACGCCACAGCAAGAGGCGCTCTAGGGTGCATATCAC	3476
QY	21	SerleuthrGlyArgAspLysAsnGlnValGlnGlyGlnValGlnileValSerThra	40
DB	3477	AGCTTAATGCGCCGAGCAAAACCAAGTGAAGGTGAAGTTCAGATTGTCAACTGCT	3536
QY	41	AlaGlnThrPheLeuAlaThrCysileasnGlyValCysTrpThraValYrHileGlyAla	60
DB	3537	GCCCAAACTTCTCCGCAAGCTGATCAATGGGTGTGCTGATGCTGCTACACAGGGGCC	3596
QY	61	GlyThrArgThrIleAlaSerProlysglyProValIleGlnMetYrThrAsnValAsp	80
DB	3597	GGAACGAGAGCAATCGGTGTCACCAAGGGTCTGTCTCATCCAGATGTATCAATGTAGAC	3656
QY	81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
DB	3657	CAAAACCTTGCGGTGCTGCGCCGTCGCAAGTAGCGCTCATTTAGACACCTTGACCTTCG	3716
QY	101	GlySerSerAspLeuYrLeuValThraGlnIleAlaAspValIleProValArgArgArg	120
DB	3717	GGCTCTCGGACCTTACCTGTCACAGAGCACCCCATGTCATTCCTGCGCGCGCGG	3776
QY	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLysGlySerSer	140
DB	3777	GGTATAGACAGGGGACGCTGCTGCGCCGCGCCATTCTTCACTTGAAGAGGCTCTCG	3836
QY	141	GlyGlyProleuLeuCysProAlaGlyHileAlaValGlyIlePheAlaAlaValCys	160
DB	3837	GGGGGTCCGCTGTGTGCGCGCGGACGCGGTGGCATATTAGGGCGCGCTGTGC	3896
QY	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGlnThrThreMet	180
DB	3897	ACCGGTGAGTGGCTTAAGCGGTGACTTTATCCCTTGAGAACTTAAGAGACACCATG	3956
QY	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal	200
DB	3957	AGGTCCCGGAGTTCACGAGTAATCTCTCCACAGAGTGGCCCAAGAGCTTCAGGTG	4016
QY	201	AlaHileuHileAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
DB	4017	GCTCACTTCATGCTCCACAGGACGCGCAAAAGCACAAAGTCCCGGCTGCATATGCA	4076
QY	221	AlaGlnGlyYrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
DB	4077	GCTCAGGAGCTTAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACAGGGCTTTGCT	4136
QY	241	AlaTyrMetSerLysAlaHileGlyLysAspProAsnIleArgThrGlyValArgThrIle	260
DB	4137	GCTTAATGTCCAGGCTCATGGATGATCTTAATCATCAGACCGGGGTGAAGAACAATT	4196
QY	261	ThrThrGlySerProIleThrYrSerThrYrGlyLysPheLeuAlaAspGlyGlyCys	280
DB	4197	ACCACTGGAGCCCATCACTACTCCACCTTAAGCAAGTTCCTTTGCGAGGGGGTGC	4256

QY	281	SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHileSerThraAlaThrSer	300
DB	4257	TCGGGGGGCGCTTATGACATTAATTTGAGAGAGGCCACTCCACGATGCCATCC	4316
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaGlyLeuThrVal	320
DB	4317	ATCTGGGCAATCGCACTGCTCTTGAACAGAGAGACTGCGGGGCGAATCGTTGTG	4376
QY	321	LeuAlaThrAlaThrProProGlySerValThraValProHileProAsnIleGlnVal	340
DB	4377	CTGCGCACCGCACCCCTCGGGCTCGTACTGTGCCCATCCCAATCGAGAGGATT	4436
QY	341	AlaLeuSerThrThrGlyGlnIleProPheYrGlyLysAlaIleProLeuGlnAlaIle	360
DB	4437	GCTGTGCAACACCGAGAGATCCCTTTTACGGAAGGCTATCCCTCCGAAGTATC	4496
QY	361	LysGlyGlyArgHileuHilePheCysHileSerLysLysLysCysAspGlyLeuAlaAla	380
DB	4497	AAAGGGGGGAGATCTCATCTTCTGTCAAAAGAGAGGCAACGACTCGCGCA	4556
QY	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
DB	4557	AACTGTGCTCATTTGGGCAATGCGCGCTTACTACCGCGCTTGAAGTGTCCGCTC	4616
QY	401	IleProThrSerGlyAspValValAlaAlaThrAspAlaLeuMetThrGlyPheThr	420
DB	4617	ATCCGACAGCGGGCATGTGTGTGTGTGGCAACGATCCCTCATGACGGCTATAC	4676
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
DB	4677	GGGACCTTCACCTCGGAGTAGCTCAATACCTGTGTCACCCAGACAGTGCATTCGAC	4736
QY	441	LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr	460
DB	4737	CTTGACCTTACCTTACCATTAAGACATACCTCCCGAGATGCTGTCCGCACT	4796
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
DB	4797	CAAGCTCGGGGAGAGCTGGCAAGGGGAAAGCCAGCATCTACAGATTGTGGACCGGG	4856
QY	481	GlnArgProSerGlyMetPheAspSerSerValLeuCysGlyCysTyrAspAlaGlyCys	500
DB	4857	GAGCGCCCTCCGGCATGTGTGCACTGCTGCTGTGAGAGTATGACGAGGCTGT	4916
QY	501	AlaThrProGlyLeuThrProAlaGlyIleThrThrValArgLeuArgAlaTyrMetAsnThr	520
DB	4917	GCTTGATATGAGTCAAGCCCGCGAGACTACGTTAGGCTACAGAGCATGAAACCC	4976
QY	521	ProGlyLeuProValCysGlnAspHileuGlnPheTrpGlyValPheThrGlyLeu	540
DB	4977	CCGGGCTTCCCTGTGCGCAGAGCATCTTGAATTTGGAGGGCGCTTTACAGGCTC	5036
QY	541	ThrHileIleAspAlaHilePheLeuSerGlnThrLysGlnSerGlyLysLeuProYr	560
DB	5037	ACTCATATGATGCCCATCTTATCCAGACAAACAGAGTGGGAGAACTTCTCTTAC	5096
QY	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp	580
DB	5097	CTGTATGCTTCAACCAACCCAGTGTGCTAGGGCTCAAGCCCTCCCATGCTGGAG	5156
QY	581	GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHileGlyProThrProLeuLeu	600
DB	5157	CAGATGTGAAGTGTGATGCTCCACAGCCACCTTCATGGGCAACACCCCTGCTA	5216
QY	601	TyrArgLeuGlyAlaValGlnAsnGlnValThrLeuThrHileProValThrLysTyrIle	620
DB	5217	TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGACCACTCAACCAATATCATC	5276
QY	621	MetThrCysMetSerLysLeuGlnValValThrSerThrTyrValLeuValGlyGly	640
DB	5277	ATGACATGATGTGCGCGCACTGAGGTGTGACAGACCTGGGTGCTGTGGCGGG	5336
QY	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660

```

Db      5337  GTCTGGCTGCTTGGCGCGCTATTCCTGCATCAACAGCGCTGCGTCATATGTCGCGCAGG 5396
Qy      661  ILleValIeuSerGlySerProAlaIleIleProAspArgGluValIleuTyrArgGluPhe 680
Db      5397  GTCTGCTTGTGCGGAGAACCGGCAATCACTTAACAGGAGAGTCTTACCGAGAGTTCC 5456
Qy      681  AspGluMetGluGluCys 686
Db      5457  GATGAGATGGAAGAGTGC 5474

RESULT 14
AAD35043
ID      AAD35043 standard; cDNA; 9401 BP.
XX
XX      AAD35043;
XX
XX      16-JUL-2002 (first entry)
XX
XX      Hepatitis C virus (HCV) polypeptide cDNA.
XX
XX      Hepatitis C virus; HCV; antigen; C domain; polypeptide; NS3 domain;
XX      NS4 domain; S domain; NS5 domain; gene; ss.
XX
XX      Hepatitis C virus.
XX
XX      Key
XX      CDS      Location/Qualifiers
XX              342..9377
XX              /*tag= a
XX              /product= "HCV polypeptide"
XX              /transl_except= (pos:366..368, aa:Xaa)
XX              /note= "Xaa equals Lys or Arg"
XX              /transl_except= (pos:372..374, aa:Xaa)
XX              /note= "Xaa equals Thr or Asn"
XX              /transl_except= (pos:867..869, aa:Xaa)
XX              /note= "Xaa equals Thr or Ile"
XX              /transl_except= (pos:1341..1343, aa:Xaa)
XX              /note= "Xaa equals Val or Met"
XX              /transl_except= (pos:2148..2150, aa:Xaa)
XX              /note= "Xaa equals Ile or Leu"
XX              /transl_except= (pos:2883..2885, aa:Xaa)
XX              /note= "Xaa equals Asn or Tyr"
XX              /transl_except= (pos:3681..3683, aa:Xaa)
XX              /note= "Xaa equals Ser or Pro"
XX              /transl_except= (pos:3690..3692, aa:Xaa)
XX              /note= "Xaa equals Thr or Ser"
XX              /transl_except= (pos:4167..4169, aa:Xaa)
XX              /note= "Xaa equals Leu or Pro"
XX              /transl_except= (pos:4323..4325, aa:Xaa)
XX              /note= "Xaa equals Val or Gly"
XX              /transl_except= (pos:4701..4703, aa:Xaa)
XX              /note= "Xaa equals Tyr or Cys"
XX              /transl_except= (pos:4752..4754, aa:Xaa)
XX              /note= "Xaa equals Ser or Thr"
XX              /transl_except= (pos:5970..5972, aa:Xaa)
XX              /note= "Xaa equals Gly or Glu"
XX              /transl_except= (pos:6183..6185, aa:Xaa)
XX              /note= "Xaa equals His or Leu"
XX              /transl_except= (pos:6186..6188, aa:Xaa)
XX              /note= "Xaa equals Cys or Ser"
XX              /transl_except= (pos:6402..6404, aa:Xaa)
XX              /note= "Xaa equals Gly or Val"
XX              /transl_except= (pos:7386..7388, aa:Xaa)
XX              /note= "Xaa equals Ser or Thr"
XX              /transl_except= (pos:7494..7496, aa:Xaa)
XX              /note= "Xaa equals Phe or Tyr"
XX              /transl_except= (pos:7497..7499, aa:Xaa)
XX              /note= "Xaa equals Ala or Ser"
XX              /transl_except= (pos:7845..7847, aa:Xaa)
XX              /note= "Xaa equals Phe or Leu"
XX              /transl_except= (pos:8409..8411, aa:Xaa)
XX              /note= "Xaa equals Gly or Arg"
XX              /transl_except= (pos:9102..9104, aa:Xaa)

```

```

FT      FT      /note= "Xaa equals Gly or Arg"
FT      FT      /transl_except= (pos:9327..9329, aa:Xaa)
FT      FT      /note= "Xaa equals pro or leu; these translational
FT      FT      exceptions occur while decoding for the alternative
FT      FT      version of HCV polypeptide (AAE22052)"
XX      XX      US6312889-B1.
XX      XX      06-NOV-2001.
XX      XX      12-MAY-1995; 95US-00440549.
XX      XX      04-APR-1990; 90US-00504352.
XX      XX      07-JUL-1992; 92US-00910760.
XX      XX      (CHIR ) CHIRON CORP.
XX      XX      Houghton M, Choo Q, Kuo G;
XX      XX      WPI; 2002-040268/05.
XX      XX      P-PSDB; AAE22049.
XX      XX      Combination of hepatitis C viral (HCV) antigens, useful in improved
XX      XX      immunoassay for detecting HCV antibodies.
XX      XX      Example 1; Fig 1; 58pp; English.
XX      XX      The invention relates to combination of hepatitis C viral (HCV) antigens
XX      XX      that have a broader range of immunological reactivity than any single HCV
XX      XX      antigen. The combinations consist of an antigen from the C domain of the
XX      XX      HCV polypeptide, and at least one additional HCV antigen from either the
XX      XX      NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
XX      XX      the form of fusion protein, a simple physical mixture, or the individual
XX      XX      antigens commonly bound to a solid matrix. The combinations of antigens
XX      XX      provides broad range immunoassays for anti-HCV antibodies. The invention
XX      XX      therefore provides a method for detecting antibodies to HCV in a mammal
XX      XX      suspected of containing such antibodies. The present sequence is HCV
XX      XX      polypeptide cDNA
XX      XX      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
XX      XX      Alignment Scores:
XX      XX      Pred. No.: 1,936-241 Length: 9401
XX      XX      Score: 3574.00 Matches: 672
XX      XX      Percent Similarity: 99.42% Conservative: 10
XX      XX      Best Local Similarity: 97.96% Mismatches: 4
XX      XX      Query Match: 98.78% Indels: 0
XX      XX      DB: 6 Gaps: 0
XX      XX      US-09-930-591-2 (1-686) x AAD35043 (1-9401)
Qy      Qy      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      Db      3417 CTGGCGCCATCAACGGGTACCGCCAGCAAGAGGGGCTCTTAAGGTGCATATACACC 3476
Qy      Qy      21 SerIeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      Db      3477 AGCTTAACGTGGCGGCAAAAACCAAGTGAAGGGTGAAGTTCAGATTGTGTAACGTCT 3536
Qy      Qy      41 AlAGlnThrPheIleuAlaThrCysIleAsnGlyValCysThrPheValTyrHisGlyAla 60
Db      Db      3537 GCCCAAACTTCTCGGACAGTCATCAATGAGGTGGTGGAGCTGTCTACACAGGGGCC 3596
Qy      Qy      61 GltYhrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      Db      3597 GGAACGAGACCAATCGGTACCCAGAGGCTCTGTATCAGATGTATACCAATGTATGAC 3656
Qy      Qy      81 GlnAspIeuValGlyTyrProAlaProGlnGlyAlaArgSerIeuThrProCysGlnCys 100
Db      Db      3657 CAAGACCTTGTGGGTGGCGCGCTCGCAAGGTAGCGCGTCATGTGACACCTTCACATTGC 3716
Qy      Qy      101 GltSerSerIeuIeuValThrArgHisAlaAspValIleProValArgArgArg 120

```

Db 3717 GGCCTTCGACCTTAACTGCTGACAGAGCAGCCCATGTCATTCCCGTCCGCGCGG 3776
 QY 121 G1YAspG1YArG1YSerLeuLeuSerProArProIleSerTYrLeuYsG1YSerSer 140
 Db 3777 GGTGATAGCAGAGGCGACCTGCTGTGCGCCCGCCATTTCCTAATTGAAAGGCTCTCG 3836
 QY 141 G1YG1YProLeuLeuCySProAlaG1YH1Sa1aValG1Y1IePheArG1a1a1aValCys 160
 Db 3837 GGGGGTCCGCTGTGTGTCGCCCGCGGCGACGCCGTGGGCAATATTAAAGGCGCGCTGTGC 3896
 QY 161 ThrArG1YVal1a1a1aValAspPhe1IeProValG1Guse1LeuG1uThrThMet 180
 Db 3897 ACCGTGAAGTGGCTAAAGCGGTGACTTAACTCTTGGAAGAACTTAAGACACCAATG 3956
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProG1nSerTYrG1nVal 200
 Db 3957 AGGTCCCGGTGTTTACAGGATTAATCTCTCCACCAAGTAAAGTCCCGCTGCATATGCA 4016
 QY 201 AlaH1SleuH1Sa1aProThrG1YSerG1YLySserThrLySValProAla1a1aTYrAla 220
 Db 4017 GCTCACTCCATGCTCCACAGCGAGCGGCAAAAGCAACAAAGTCCCGCTGCATATGCA 4076
 QY 221 AlaG1nG1YTYrLySValLeuValLeuAsnProSerVal1a1a1aThrMetG1YPhG1Y 240
 Db 4077 GCTCAGGCTAATAGGTGCTAGTACTCAACCCCTCTGTGTGCAACACAGGCTTTGTG 4136
 QY 241 AlaTYrMetSerLySa1aH1Sg1Y1IeAspProAsn1IeArG1ThrG1YValArgThr1Ie 260
 Db 4137 GCTTACATGTCGAAGGCTCATGGATGATCTTAACATCAGAGCCGGGTGAGAACAAATT 4196
 QY 261 ThrThrG1YSerProIleThrTYrSerThrTYrG1YLySPhLeuAlaAspG1YG1YCys 280
 Db 4197 ACCATGGCAGACCCCATCACTACTCCACCTCAAGGATTCCTTGCAGAGCGGCTGTGC 4256
 QY 281 SerG1YG1YValATYrAsp1Ie1Ie1IeCyAspG1YCySH1SserThrAspAlaThrSer 300
 Db 4257 TCGGGGGGCGCTTAATGACATTAATTGTGACAGAGTGCACCTCACAGGATGCCACATCC 4316
 QY 301 1IeLeuG1Y1IeG1YThrValLeuAspG1n1aG1uThr1aG1YAlaArgLeuThrVal 320
 Db 4317 ATCTTGAGCATCGGACCTGCTCTTGACCAAGACAGACTGGGGGGGAGACTGTGTGTG 4376
 QY 321 LeuAlaThrAlaThrProProG1YSerValThrValProH1SProAsn1IeG1uG1YVal 340
 Db 4377 CTGGCCACCGCCACCCCTCCGGGCTCCGTACCTGTGCCCATCCCAATCGAGAGATT 4436
 QY 341 AlaLeuSerThrThrG1YG1u1IeProPheTYrG1YSa1a1IeProLeuG1uAla1Ie 360
 Db 4437 GCTCTGTCCACCAACCGAGAGATCCCTTTTAAACGCAAGGCTATCCCTCGAAGTATC 4496
 QY 361 LySg1Yg1YArG1H1Sleu1IePheCySH1SserLyS1YsCyAspG1uLeuAla1a 380
 Db 4497 AAGGGGGAGACATCTCATCTCTGTCAATTCAAGAGAGTGCAGACGAACTGCCGCA 4556
 QY 381 LySLeuVal1a1aLeuG1YValAsnAlaVal1a1aTYrTYrArgG1YLeuAspValSerVal 400
 Db 4557 AAGCTGTCGATTTGGGCAATAAGCCGTGCCCTACACCGCGCTCTTGAAGTTCGCTC 4616
 QY 401 1IeProThrSerG1YAspValVal1a1aVal1a1aThrAspAlaLeuMetThrG1YPhThr 420
 Db 4617 ATCCCGACCAAGCGCGAGTGTGTGTGTGTCGAACCATGCTCCATGACCGGCTAATAC 4676
 QY 421 G1YAspPheAspSerVal1IeAspCyAsnThrCySValThrG1uThrValAspPheSer 440
 Db 4677 GGCACCTTCGACTGGGATAGACTGCATAACGTGTGTGCCACCCAGACAGTGCATTTACGC 4736
 QY 441 LeuAspProThrPheThr1IeG1uThr1IeThrLeuProG1nAspAlaValSerArgThr 460
 Db 4737 CTTGACCTCACTTCAACATTAAGACATACGCTCCCGACAGATGTGTCTCCGCACT 4796
 QY 461 G1nArgArG1YArG1ThrG1YArG1YLySProG1Y1IeTYrArgPheValAlaProG1Y 480
 Db 4797 CAACTCGGGGCAAGACTGGCAGGGGGAAGCCAGGCAATTAACAATTTGTGGACCGGGG 4856

QY 481 G1nArgProSerG1YMetPheAspSerSerValLeuCySg1YCyTYrAspAlaG1YCyS 500
 Db 4857 GAGCGCCCTCCCGCATGTTGCACTGCTCGTCTGTGAAGTATGACGAGGCTGT 4916
 QY 501 AlaTYrTYrG1uLeuThrProAlaG1uThrThrValArgLeuArg1a1aTYrMetLeuThr 520
 Db 4917 GCTTGATATGAGTCAACCGCCCGAGACTACAGTTAAGCTTACAGAGTACGAACACC 4976
 QY 521 ProG1YLeuProValCySg1nAspH1SleuG1uPheTYrG1uG1YVal1PheThrG1YLeu 540
 Db 4977 CCGGGCTTCCCTGTGTGCCAGACCATCTTGATTTTGGAGGGCGCTTTACAGGCTTC 5036
 QY 541 ThrH1S1IeAspAlaH1SPhLeuSerG1nThrLySg1nSerG1Yg1uLeuLeuProTYr 560
 Db 5037 ACTCATATATGATGCCCATCTTATCCAGACAACAAGAGTGGGAGAACTTCTCTTAC 5096
 QY 561 LeuValAlaTYrG1nAlaThrValCYBa1aArgAlaG1nAlaProProProSerTPAsp 580
 Db 5097 CTGGTAGCGTACCAAGCCACCGGTGCGCTAAGGGCTCAAGCCCTCCCATGCTGGAGC 5156
 QY 581 G1nMetTYrLySAsp1IeArG1LeuLySProThrLeuH1Sg1YProThrProLeuLeu 600
 Db 5157 CAGATGTGAAGTGTGATTCCCTCAAGCCCATCCATGGGCCAACACCCCTGCTCA 5216
 QY 601 TYrArgLeuG1YAlaValG1nAsnG1uValThrLeuThrH1SProValThrLyS1Y1Ie 620
 Db 5217 TACAGACTGGCGCTGTGAGATGAATATCACTCAGCAGCACCAGTCAACAAATACATC 5276
 QY 621 MetThCYrMetSerAlaAspLeuG1uVal1a1aThrSerThrTYrValLeuValG1YArg 640
 Db 5277 ATGACATGCAATGTGCGCCGACCTGAGGTGCTGACAGCACCCTGGGTGCTGTGGCGGC 5336
 QY 641 ValLeuAla1a1aLeuAla1a1aTYrCySLeuSerThrG1YCySVal1a1IeVal1G1YArg 660
 Db 5337 GTCTGTGCTGTTTGGCGGATGCTGTCTGTCAACAGGCTGCTGTGATAGTGGGAGG 5396
 QY 661 1IeValLeuSerG1YLySProAla1Ie1IeProAspArgG1uVal1euryArgG1uPhe 680
 Db 5397 GTGCTGTGTGCGGGAAGCGGCAATCATATCCGACAGGGAATCCCTCAACGAGTTCC 5456
 QY 681 AspG1uMetG1uG1YCyS 686
 Db 5457 GATGAGATGGAAGATGTC 5474

RESULT 15
 AAN92103 standard; DNA; 6905 BP.
 ID AAN92103
 AC AAN92103;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX
 DE Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones 12f through 15e.
 XX
 KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..6905
 FT /*tag= a
 XX
 PN EP318216-A.
 XX
 PD 31-MAY-1989.
 XX
 PF 18-NOV-1988; 88BP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.
 PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 PA (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo QL, Kuo G;
 XX
 DR WPI, 1989-159274/22.
 DR P-PSDB; AAP92047.
 XX
 PT Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX
 PS Claim 3; Fig 32-1 - 32-7; 139pp; English.
 XX
 CC It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2.16e-241 Length: 6905
 Score: 3571.00 Matches: 671
 Percent Similarity: 99.42% Conservative: 11
 Best Local Similarity: 97.81% Mismatches: 4
 Query Match: 98.70% Indels: 0
 Gaps: 0
 US-09-930-591-2 (1-686) x AAN92103 (1-6905)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1303 CTGCGCCCATCATCGCGCTACGCCAGACGACGAGGCGCTCTGAGGTCATATATCAC 1262
 QY 21 SerLeuThrGlyArgAspGlyAsnGlnValGlnGlyGlnValGlnIleValSerThrAla 40
 DB 1263 AGCTTAAGTGGCCGGGCAAAACCAAGTAGAGGTAGATCCAGATTGTCTCACTGCT 1322
 QY 41 AlaGlnThrPheLeuAlaIleThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 1323 GCCCAAACTTCCTGCGCAACCTGCATCATATGGGTGTGCTGACTGTCTACCAAGGGGCC 1382
 QY 61 GlyThrArgThrIleAlaSerProGlySerGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 1383 GGAACGAGACCATCGCTGCTACCCAAAGGCTCTGTCTCATCGATGATATACCAATGTAGAC 1442
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 1443 CAAGACCTTGTGGGCTGGCCGGCTCCGCAAGTAGCCGCTCATTTGACACCTGCATTTGC 1502
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 1503 GGCCTCTCGGACCTTTACTGTGTACGACGACGACCCCGATGTCAATTCCTGTCGCGCGG 1562
 QY 121 GlnAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
 DB 1563 GGTATATGACAGGGGAGGCTGTGTGTGCGCCCGCCCATTTCTATCTTGAAGAGGCTCTCG 1622
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 1623 GGGGGTCCGCTGTGTGTGCGCGCGGAGCGCCGTGGGCATATTTAAGGCGCGGTGTGC 1682
 QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

DB 1683 ACCCGTAGAGTGGCTAAGCGGTGACCTTATCTCTGTGGAGAACCTAGACACCATG 1742
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 1743 AGGTCGCCGCTGTTCACGAGTAACTCTCTTCCACAGTAACTGCCCCAGACTTCCAGGTG 1802
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyValSerThrIleValProAlaAlaTyrAla 220
 DB 1803 GCTCACTCATGCTCTCCACAGGACGCGGCAAAAGCACCAAGGATCCCGGCTGCATATGCA 1862
 QY 221 AlaGlnGlyTyrTyrValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 1863 GCTCAGGCTATAGGTGCTAGTACCAACCTCTGTGTCTGCAACACTGGGCTTTGGT 1922
 QY 241 AlaTyrMetSerSerValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 1923 GCTTACATGTCGAAGGCTCATGGGATCGATCTTAACATCGAGCCGGGGTGAACAATT 1982
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyValPheLeuAlaAspGlyGlyCys 280
 DB 1983 ACCACTGGACGCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2042
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 2043 TCGGGGGGCGCTTATGACATATATTTGTGAGACTGCCACTCCACGATGCCACATCC 2102
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
 DB 2103 ATCTTGGCATTCGACATCTGTCTTGAACAGACAGACTCGGGGGGAGACTGTGTGTG 2162
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
 DB 2163 CTGCGCACCGCCACCCCTCGGGCTCCGTCACTGTCACATGCCCATCCAACTCAGAGGTT 2222
 QY 341 AlaLeuSerThrThrGlyGlyIleProPheTyrGlyValAlaIleProLeuGlnAlaIle 360
 DB 2223 GCTCTGTCCACCAACGAGAGATCCCTTTTAAAGGAGCATATCCCTCGAAGTAAATC 2282
 QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerLeuLysHisCysAspGlnLeuAlaAla 380
 DB 2283 AAGGGGGGAGACATCTCATCTTCTGTCTTCAAAAGAGAGTCCAGACATCTCCGCGCA 2342
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 2343 AAGCTGGTCGATTGGGCATCATGCGGTGCTTACCTTCAACGCGGTCTTGAACGTCTC 2402
 QY 401 IleProThrSerGlyAspValValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 2403 ATCCGACCAAGCGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2462
 QY 421 GlnAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 2463 GCGGACTTGCATCTCGGTGATGACTGCATACGTGTGTGTGTGTGTGTGTGTGTGTGT 2522
 QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 2523 CTTCACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2582
 QY 461 GlnArgArgGlyArgThrGlyArgGlyValProGlyIleTyrArgPheValAlaProGly 480
 DB 2583 CAACGTTCGGGAGGAGCTGGCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2642
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyGluCysTyrAspAlaGlyCys 500
 DB 2643 GAGCGCCCTCTCGGCATGTTCACATCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2702
 QY 501 AlaTrpTyrGlnLeuThrProAlaGlnThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 2703 GCTTGTGATAGCTCAAGCGCCGAGAGCTACAGTTAAGCTACAGACGTACAGAACAC 2762
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheThrGlnGlyValPheThrGlyLeu 540
 DB 2763 CCGGGGCTTCCCGT 2822

```
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyIleAsnLeuProTyr 560
    |||||
Db 2823 ACTCATATAGATGCCCACTTCTATATCCAGACAAAGCAGAGTGGGAGAACTTCCCTTAC 2882
    |||||
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPseTTPAsp 580
    |||||
Db 2883 CTGGTAGCCGTACCAAGCCACCGTGTGCTCAGGGCTCAAGCCCCCTCCCATCGTGGAC 2942
    |||||
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
    |||||
Db 2943 CAGATGTGAAAGTGTGATTGATTCGCTCAAGCCACCTCCATGGGCCACACCCCTGCTA 3002
    |||||
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
    |||||
Db 3003 TACAGACTGGGCGCGCTGTCAGAATGAATCACCCCTGACGACCCCAATACATC 3062
    |||||
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyIy 640
    |||||
Db 3063 ATGACATGCATGTCTGGCCGACCTGAGGTGCTCACGACGACCTGGGTGCTGTTGGCGGC 3122
    |||||
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
    |||||
Db 3123 GTCTGGCTGCTTGGCGCGATATGCTGTCAACAGGCTGCGTGCATATAGTGGCAGG 3182
    |||||
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
    |||||
Db 3183 GTCGTCTTGTCCGGAAAGCCGGCAATCATCTGTACAGGAAAGTCTTACCGAGAGTTTC 3242
    |||||
Qy 681 AspGluMetGluGluCys 686
    |||||
Db 3243 GATGAGATGGAAGAGTGC 3260
    |||||
```

Search completed: September 17, 2004, 08:29:52
Job time : 766 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 17, 2004, 07:46:20 ; Search time 127 Seconds
(without alignments)
2997.609 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAOCTRGILGCIIT.....PAIIPPREVLYREFDEMEEC 686

Scoring table: BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2_1/USFTO_epool/US09930591/rnat 13092004_164953_711/app query.fasta_1.839
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09930591 @CGN 1.1 56 @rnat 13092004_164953_711 -NCP=6 -ICPU=3
-NO_MMAP -LARGEOTERY -NEG_SCORES=0 -WAIT -DSBBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WASN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents NA:*
2: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	US-09-388-874-1
3	3574	98.8	9379	4	US-09-916-359-1
4	3574	98.8	9401	1	US-07-910-760-9
5	3574	98.8	9401	1	US-08-440-519-9
6	3574	98.8	9401	4	US-08-440-549-9
7	3574	98.8	9401	4	US-08-823-895A-25
8	3571	98.7	6785	3	US-08-444-818-65
9	3570	98.7	8316	3	US-08-444-818-88
10	3570	98.7	9185	3	US-08-444-818-122
11	3570	98.7	9185	3	US-08-444-818-123
12	3565	98.5	2058	4	US-09-881-654-1

13	3565	98.5	2058	4	US-09-881-654-1	Sequence 1, Appl1
14	3565	98.5	8987	3	US-08-444-818-137	Sequence 137, App
15	3555	98.3	5360	3	US-08-444-818-53	Sequence 53, Appl1
16	3551	98.1	9646	3	US-08-811-566-1	Sequence 1, Appl1
17	3551	98.1	9646	4	US-09-034-756-1	Sequence 1, Appl1
18	3551	98.1	12980	3	US-08-811-566-5	Sequence 5, Appl1
19	3551	98.1	12980	4	US-09-034-756-5	Sequence 5, Appl1
20	3544	98.0	9599	3	US-09-014-416-2	Sequence 2, Appl1
21	3544	98.0	9599	3	US-09-014-416-6	Sequence 6, Appl1
22	3538	97.8	9401	2	US-08-432-693-1	Sequence 1, Appl1
23	3538	97.8	9416	3	US-08-811-566-19	Sequence 19, Appl1
24	3538	97.8	9416	4	US-09-539-601-7	Sequence 7, Appl1
25	3535	97.7	9379	3	US-08-444-818-176	Sequence 176, App
26	3532	97.6	9401	5	PCT-US91-02225-9	Sequence 26, Appl1
27	3504	96.8	9416	4	US-08-823-895A-26	Sequence 1, Appl1
28	3504	96.8	9416	4	US-10-104-966-13	Sequence 13, Appl1
29	3429	94.8	7989	4	US-09-539-601-10	Sequence 10, Appl1
30	3429	94.8	8001	4	US-09-539-601-7	Sequence 7, Appl1
31	3426	94.7	8637	4	US-09-539-601-4	Sequence 4, Appl1
32	3426	94.7	8649	4	US-09-539-601-13	Sequence 13, Appl1
33	3426	94.7	11076	4	US-09-539-601-1	Sequence 1, Appl1
34	3423	94.6	8001	4	US-09-539-601-22	Sequence 22, Appl1
35	3420	94.5	6039	1	US-08-384-616-11	Sequence 11, Appl1
36	3420	94.5	6039	2	US-08-384-616-11	Sequence 11, Appl1
37	3420	94.5	6039	2	US-08-904-686A-11	Sequence 11, Appl1
38	3420	94.5	6039	3	US-09-315-850-11	Sequence 11, Appl1
39	3420	94.5	7863	1	US-08-324-977-35	Sequence 35, Appl1
40	3420	94.5	7863	2	US-08-384-616-35	Sequence 35, Appl1
41	3420	94.5	7863	2	US-08-904-686A-35	Sequence 35, Appl1
42	3420	94.5	7863	3	US-09-315-850-35	Sequence 35, Appl1
43	3420	94.5	7917	1	US-08-324-977-31	Sequence 31, Appl1
44	3420	94.5	7917	2	US-08-384-616-31	Sequence 31, Appl1
45	3420	94.5	7917	2	US-08-904-686A-31	Sequence 31, Appl1

ALIGNMENTS

RESULT 1
US-08-444-818-74
Sequence 74, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
ADDRESSER: Rutten, William J.
TITLE OF INVENTION: NANBY Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSER: Chilton Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444, 818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403, 590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hardin, Aisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:
LENGTH: 7310 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..7310
US-08-444-818-74

Alignment Scores:
Pred. No.: 0 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-74 (1-7310)

QY 1 MetLAbroilleThra1aTyrAlaGlnGlnThraArgLysLeuGlyCysIleIleThr 20
DB :::
DB 1728 CTGGCGCCCATCAGCGGTACGCCAGACAGAGGGCCCTCTAGGGGTGATATACACC 1787
QY 21 SerLeuThrgLysArgAspLysAsnGlnValGlnGlyLysValGlnIleValSerThra1a 40
DB AGCTTAATGCGCCGGGACAAACCAACAGTGAAGGGTGAAGTCCAGATTGTCTACACGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIYrHisGlyVala 60
DB GCCCAAACTTCTCGGCAACGTCATCATGGGGTGTGTGAGCTGTCTACACAGGGGCC 1907
QY 61 GlyThrArgmTrlIeAlaSerProLysGlyProValIleGlnMetYrThrAsnValasp 80
DB GAAACGAGACCATCGCTCACCCAGAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyValaArgSerLeuThrProCysThrCys 100
DB CAACACCTTGAGGCTGGCCCGCTCCGAGAGTACCGGCTCATTTAGACCCCTGACATTGC 2027
QY 101 GlySerSerAspLeuYrLeuValThraArgHisAlaAspValIleProValArgArgArg 120
DB GGCTCCCTCGGACCTTATCTGTGTACAGAGGACCGCCATGTCTCCCGTGGCGCGCGG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerYrLeuLysGlySerSer 140
DB GGTAATACAGAGGGGAGGCTGTGTGCGCCGCGCCCATTTCTTCTTGAAGAGGCTCTCG 2147
QY 141 GlyLysProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB GGGGGTCCGCTGTTGTGCCCCGGGGGACGCGGTGGGCATATTAGGGCGCGGTGTGC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThMet 180
DB ACCCGTGAAGTGCCTAAGCGGTGACTTATTCCTGTGAGAACCTTAAGACCAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerYrGlnVal 200
DB AGGTCCCGGTGTTCACGATTAACCTCTCTCACCAAGTGAAGTCCAGAGCTTCAAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrgLysSerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB GCTACCTCCATGTCTCCACAGGAGGAGGCAAAAGACCAAGTCCCGGGCTGATAGCA 2387
QY 221 AlaGlnGlyYrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB GCTACGAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 2447
QY 241 AlaThrMetSerLysAlaHisGlyIleAspProAsnIleArgmTrgValArgThrIle 260
DB GCTTAACATGTCCAAAGGCTCATGGATGTATCAACATCAGACCGGGGTGAGAACATTT 2507

QY 261 ThrThrgLysSerProIleThrYrSerThrYrGlyLysPheLeuAlaAspGlyGlyCys 280
DB ACCACTGGACGAGCCCATCAAGTACTCACCTACGGCAAGTTCCTTCCGACGGCGGGTGC 2567
QY 281 SerGlyGlyValaYrAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
DB TCGGGGGGGCTTATACATTAATTAATTTGTGAGAGATGTGCATCCACGATGCAACATCC 2627
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
DB ATCTTGGGCAATGGGCACTGTCTCTTACCAAGACAGACATCGGGGGCGGAGACTGTGTG 2687
QY 321 LeuAlaThrAlaThrProProGlySerValThraValProHisProAsnIleGlnIleVal 340
DB CTGCGCACCGCCACCCCTCGGGCTCGTCACTGTGCCCCCATCCCAACATCGAGAGGTT 2747
QY 341 AlaLeuSerThrThrgLysGlnIleProPheYrGlyLysAlaIleProLeuGlnAlaIle 360
DB GCTGTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATVCCCTCCGAAATTAATC 2807
QY 361 LysGlyValArgHisLeuIlePheCysHisSerLysLysLeuCysAspGlnLeuAlaAla 380
DB AAGGGGGGAGACATCTCATCTTCTGTCAATCAAAAGAAAGTGCACGAACTCGCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB AACCTGGTGCATTTGGGCAATCAATGCCGTGCTTACCCGGCGGTCTTACGTCGTCCGTC 2927
QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB ATCCCAACAGCGGCATGTGTGTGTGTGCAACGGAATGCCCTCATGACCGGCTATVACC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrgLlnThrValAspPheSer 440
DB GCGCACTTGCATCTCGGTGTATGACTGCATACGTGTGTACCCAGACAGTCAATTTCAG 3047
QY 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB CTTGACCTTACCTTCAACATTTGAGACATACACGCTCCCGAGATCTGTCTCCCGACCT 3107
QY 461 GlnArgArgGlyArgmTrgValArgGlyLysProGlyIleThrArgPheValAlaProGly 480
DB CAACGTGGGGGAGAGCTGACAGAGGGGAAAGCCAGGATCTTACAGATTGTGGACCGGG 3167
QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGlnCysYrAspAlaGlyCys 500
DB GAGCGCCCTCCCGGCATGTTTCAGCTGTCTGTGTGTGAGTGCATGACGAGGCTGT 3227
QY 501 AlaTrpYrGlnLeuThrProAlaGlnThrThraValArgLeuArgAlaTyrMetAsnThr 520
DB GCTTGGTATGAGCTCAAGCCCGCGAGACTACAGTTAAGGTACAGACGTAATGAACACC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheThrGlnGlyValPheThrGlyLeu 540
DB CCGGGGCTTCCGTGTGCGAGACATCTTGAATTTGGAGGGCGCTTTTACAGGCTTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLysAsnLeuProYr 560
DB ACTCATATATAGATCCCACTTCTTATCCAGACAAACAGAGTGGGAGAACTTCTTTC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
DB CTGGTAGCGTACCAACCCACCGGTGCTAGGGGCTCAAGCCCTCCCATCTGTGTGGAG 3467
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB CAAATGTGGAAGTGTGATTCCTCTCAAGCCCACTTCATGGGCAACCCCTGCGTA 3527
QY 601 TyrArgLeuGlyValAlaValGlnAsnGlnValThrLeuThrHisAspProValThryLysTyrIle 620
DB TACAGACTGGGCGCTGTCTCAAGATGAATACACCTTACCCCAATCACAATATCATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGlnValValThrSerThrTrpValLeuValGlyGly 640

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCyGlnCysTyrAspAlaGlyCys 500
DB 4835 GAAGGCCCCCTCCGCAATGTTTGCATCGTCCTCTGTGAGGCTATGACGAGGCTGT 4894
QY 501 AlaTyrPylGluLeuThrProAlaGluThrThValArgLeuArgAlaTyrMetAsnThr 520
DB 4895 GCTTGGATAGCTCAGCCGCCGAGACTAGATTAGCTACGAGCGTACATGAAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPylGluValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCTCCGTGTCACGACCATCTTGAATTTTGGGAGGGCGCTTTACAGGCTTC 5014
QY 541 ThrHisLeaAspAlaHisPheLeuSerGlnThrThGlnSerGlyValuAsnLeuProTyr 560
DB 5015 ACTCATATGATGCGCCACTTTCATCCAGACAAAGCAGAGTGGGAGAACTTCTCTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrPylPasp 580
DB 5075 CTGGTACGTAACCAAGCCAGTGTGGCTAGGGCTCAAGCCCTCCCATGTCGGGAC 5134
QY 581 GlnMetTyrPylCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGCGAAGTGTGTTGATTCGCTCAAGCCCACTCCATGGGCGCAACACCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrThGlyTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTCAAGATGAATACCTCGACGACCCGATGACCAATATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATCATGTGCGCGCACTGGAAGTCTGTCACGACCACTGGGTGCTGTTGGCGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCCGCTGCTTGGCCGCGATGCTGTCAACAGCTCGGTGTCATAGTGGGACG 5374
QY 661 IleValIleAsnSerGlyLysProAlaIleIleProAspArgGluValIleTyrArgGluPhe 680
DB 5375 GTCGCTTGTCCGGGAAGCCGCAATCATCTGACAGGAAAGTCTCTACGAGAGTTC 5434
QY 681 AspGluMetGluGluCys 686
DB 5435 GATGAGATGGAAGAGTGC 5452
RESULT 3
US-09-916-359-1
; Sequence 1, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; FILE REFERENCE: PWC97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 9379
; TYPE: DNA
; ORGANISM: Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)...(9352)
US-09-916-359-1
Alignment Scores: 0 Length: 9379
Pred. No.:

Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conserved: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0
US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)
QY 1 MetaIaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3395 CTGGGCGCCATCACGCGCTTAGCCAGCAGCAAGGGGCTCTTAGGGTGCATATACACC 3454
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 3455 AGCTTAACCTGGCCGGGCAAAAACCAAGTGAAGGTGAGTCCAGATGTGTCACTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 3515 GCCCAACCTTCTGGCAACCTGATCAATGAGGTGTCTGAGCTGTACACGCGGCGC 3574
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3575 GGAAGAGGACATCGCGTCAACCAAGGATCCTGTCAATCAGATGTATACCAATGTAGAC 3634
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3635 CAAGACTTGTGGGTGGCGCTCGCTCGCAAGGTACCGCTCATTTGACACCTGCACTTGC 3694
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3695 GGCTCTCGGACTTTACTGTCACGAGGACCGCGATGTATCCGTGCGCGCGCG 3754
QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaGlyProIleSerTyrLeuLysGlySerSer 140
DB 3755 GGTGATAGCAGGGGCGCTGCTGTGCGCCGCGCATTTCTACTTGAAGGCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3815 GGGGGTCCGCTGTGTGCGCCCGGGGACGCGGTGGCATTTAGGGCCGCGGTGTC 3874
QY 161 ThrArgGlyValAlaAlaValAlaAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB 3875 ACCCGTGAAGTGGCTAAGGCGGTGAGCTTTATCCCTGTGAAGAATTAAGCAACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3935 AGGTCCCGGTGTTCAGGATATCTCTCTCCACGATGTGCGCCAGAGCTTCCAGGTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 3995 GCTACCTCCATGCTCCACAGGACGCGCAAAAGCACCMAAGTCCCGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 4055 GCTCAGGGCTATTAAGGTGCTAGTACCTCAACCCCTGTGTGCTCAACACGTGGCTTGTGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 4115 GCTTAACATGTCAGAGCTCATGGAATGATCTTAACATGACGCGGGGAGAAACATAT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 4175 ACCACTGGACGCCCATCAGTACTCCACTACGCGCAAGTCTTGGCGAGGGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4235 TCGGGGGCGCTTATGACATATATATTTGTGACAGAGTCCACTCCAGAGATGACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4295 ATCTTGGGCACTGGCACTGTCTTGCACAGCAGACAGTGGGGGCGAGCTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProGlyLysValThrValProHisProAsnIleGluGluVal 340

Db 4355 CTGGCCACCGCCACCCCTCGGCTCGTCACTGTGCCCATCCCAATCGAGGAGTT 4414
Qy 341 AAlaSerThrThrGlyIuileProPheTrgIylYsaIaIleProLeuGluAlaIle 360
Db 4415 GCTCTGCCACCAACGAGAGATCCCTTTTACGGCAAGCTATCCCCCTCGAAGTATTC 4474
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerIysIysCysAspGluLeuAlaIa 380
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTGATTAAGAAAGAGTGGCGAATCGCGCGCA 4534
Qy 381 LysLeuValAlaLeuGlyValaLeuAlaValAlaTyrTrpArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGCTCCATTTGGCATCAATGCGGTGGCTTACTACCGCGTCTTGACGTGCTCGTC 4594
Qy 401 ILProThrSerGlyAspValValaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCGACCAAGCGGCAAGTGTGCTGTGGCAACGATGCCCTCATGACCGGCTATACC 4654
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValaAspPheSer 440
Db 4655 GCGGACTTGACCTCGGTGATGACTGCAATACGTGTGTCACCCAGACAGTCGATTCAGC 4714
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTGACCTTACCTTCAACATTTAGACAAATCACGCTCCCGAGATGCTGTCCCGCAGCT 4774
Qy 461 GlnArgArgGlyArgTrpArgGlyLysArgGlyLysProGlyLysLeuArgPheValAlaProGly 480
Db 4775 CAACGTCGGGGGAGAGCTGGCAAGGGGAGACCGAGCATCTACAGATTTGGGCAACCGGG 4834
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGGCCCTCCGSCAATGTTGACTGCTCGTCTCTGTAGAGCTATGACGCAAGGCTGT 4894
Qy 501 AlaTrpTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGATGAGCTCACGCGCCGAGACATCAAGTTAGCTACGAGGCTACATGAACACC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValaPheThrGlyLeu 540
Db 4955 CCGGGGCTTCCCGTGGCCAGACCACTTGAATTTGGAGGGCGCTTTTACAGGGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleSerGlnSerGlyLeuLeuProTyr 560
Db 5015 ACTCATATATATATCCCATTTCTATCCCAAGAAAGAGTGGGAGAACTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGGTAGCGTACCAAGCAACCGTGTGGCTAGGGCTCAAGCCCTCCCATCGTGGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGAAGGTTTGTGATTCCTTCACCCACCTTCATGGGCAACACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyValaValaGlnAsnGluValaThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGGCGCTGTTCAAGATGAATCACTTCAGACCAACCAAGTCAACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValaValaThrSerThrTrpValLeuValaGly 640
Db 5255 ATGACATGATGTCGCGCGCACTGAGAGTGTCTACAGACGACTGGGTGCTCGTTGGGGGC 5314
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValaGlyArg 660
Db 5315 GTCTGTGGCTGCTTGGCGGGTATGCTGTCAACAGGCTGCGGTGTCATATAGTGGGAGG 5374
Qy 661 ILValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGTCTGTCCGGAAAGCCGCAATCATCTGACAGGAAAGTCTCTACCGAGATTC 5434
Qy 681 AspGluMetGluGluCys 686
|||||

Db 5435 GATGAGATGAGAGTGC 5452
RESULT 4
US-07-910-760-9
; Sequence 9, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366-
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372-
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867-
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341-
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2148-
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ile."
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."

FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-07-910-760-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuGlyCysIleIleThr 20
DB 3417 CTGGCCCATCAACGGCGTACGCCCAAGAGAGGGGCTCTCTAGGCTGATATCACC 3476
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCCTAATGCGCCGAGCAAAACCAAGAGGAGGTGAGTCAATGTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAACCTTCTGGCAACGTGATCATATGGGTGTCTGACGTACACGAGGCGCC 3596
QY 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGAACGAGGACATCGCGTCAACCAAGAGGTCTGTATCAATGATATACATATGAGC 3656
QY 81 GlnAspLeuValGlyTyrProAlaPProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTGTGTGGCTGGCCGCTCCGCAAGTAGCCGCTCATTTGACACCTGACCTTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3717 GGCTCCCTGGACCTTACCTTACCTGTACAGAGCAGCCGATGTATCCCGTGGCCGCGG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuIysGlySerSer 140
DB 3777 GGTATAGCAGAGGGGAGCTGTGTGCGCCCGCCCAATTTCTACTTGAAGGCTCTCTGC 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGTCGCGTGTGTGTCGCCCGCGGAGCAGCCGTGGCATATTAGGGCGCGGTGTGC 3896
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCGTGAGAGGTGTAGAGCGGTGACCTTATCTCTGTGAGACCTTAGAGACACCATG 3956
QY 181 ArgSerProValPheSerAspAspAspSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGGTCCCGGTGTACAGATACCTCTCTCCACCAATAGTCCCGAGACTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrIysValProAlaAlaTyrAla 220
DB 4017 GCTACCTCCATGCTCCACAGGAGGCAAAAGACCAAGAGTCCCGGTGATATGCA 4076
QY 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240


```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5370
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402

```

```

OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-519-9

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
Gaps: 0
US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MeclAProtllethrlAlaYrAlagInlThrArgGlyLeuLeuGlyCysIlelelthrn 20
Db 3417 CTGGGCCCATCAACGGGCTACGCCCAACAGAGGGCTCTTAGGGTGCATATACAC 3476
QY 21 SerleuthrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCTTACTGGCCGGGCAAAAACCAAGTGAAGGTGATGTCAGATTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaIleThrCysIleAsnGlyValCysThrThrValIleValIleVal 60
Db 3537 GCCCAAACTTCCCGGCAACGTCATCAAAAGGGGTGGCTGAGCTGCTACACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnIleThrThrAsnValAsp 80
Db 3597 GGAACGAGACCATCGCTACCCAAAGGTCCTCTCATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGGGGTGGCCCGCTCCGCAAGGTAGCGGCTCATTGACACCCCTGACTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
Db 3717 GGCTCTCGAACCCTTACTGTGTCAGGAGGACGCGCATGTGATTCCTCGGCGGCGG 3776

```


QY	121	GIYAPRGYIARGIISerLeuLeuSerProArgProIleSerTyrLeuysylSerSer	140
Db	3777	GGTATTAAGAGGGGCAAGCTGTGCGCCCGGCAATTTCTTAATTGAAGCTCTCG	3833
QY	141	GIYGIYProLeuLeuysProIlaGlyIleAlaValIlePhePheGlnAlaValCysI	160
Db	3837	GGGGGTCCGCTGTGTGGCCCCCGGGGCAAGCGGATGAGATTTAAGGCCCGCGGTGTC	3891
QY	161	THRAAGYIValAlaValAlaValAlaPheIleIleProValGluSerLeuGluThrThrMet	180
Db	3897	ACCCGTGAGTGGCTTAAGCGGTGGACTTTATTCCTGTGGAAACCTTAAGACAACCAAG	3956
QY	181	ARGSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	3957	AGGTCCCGGTGTTCAAGATTAACCTCTTCCACCAAGTAAGTCCCAAGCTTCCAGGTG	4011
QY	201	ALAAlaIleuAlaAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	4017	GCTCACTTCATGCTCCCAAGGACGGCAAAAGACCAAGAGTCCGGCTGCATATGCA	4076
QY	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	4077	GCTCAGGGCTTAAGGTGCTAGTACTCAACCCCTGTGTGGCAACCTGGGCTTTGGT	4136
QY	241	AlaTyrMetSerLysAlaAlaGlyIleAspProAsnIleArgThrGlyValAlaArgThrIle	260
Db	4137	GCTTACATGTCACAGGCTCATGGAGTCATCTTAACATGAGACGGGGGTGAAACAAT	4196
QY	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys	280
Db	4197	ACCACGTGGAGGCCATACGTAACCTCACTAGGAAATGCTCTTGCCGACGGCGGGTCC	4255
QY	281	SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	4257	TGGGGGGGGCTTAATGACATAATATTGTGACGAGTCCACATCCACGAGATCCACATCC	4316
QY	301	IleLeuGlyIleGlyThrValLeuAspGlnIleGluThrAlaGlyAlaArgLeuThrVal	320
Db	4317	ATCTTGGGATGGGCACTGTCTCTTGAACAAGAGACTCGGGGGCGAGCTGGTTGG	4376
QY	321	LeuAlaThrAlaThrProProGlySerValIleThrValProHisProAsnIleGluGluVal	340
Db	4377	CTCGGCAACGCCACCCCTCCGGGCTCCGTCATGTGCCCCCACTCAAGTCAAGAGGTT	4436
QY	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	4437	GCTGTGTCCACACCGGAGAGATCCCTTTTACGGGAAGGATATCCCTCGAAGTAATC	4496
QY	361	LysGlyGlyAlaArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	4497	AAGGGGGGGAGACATCTCATCTTCTGTCAATTCAAAAGAAATGGGACCAACTCCCGCA	4556
QY	381	LysLeuValAlaIleuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	4557	AAAGCTGTGTCATTGGGCAATCAATGCCGTGGCTTACTACCGGCGTCTTGAAGTGTGTC	4616
QY	401	IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr	420
Db	4617	ATCCGACCAAGCGGAGATGTTGTGTCGCGGACCAAGCAAGGCCCTCATGACGGGCTAATCC	4676
QY	421	GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrValAlaAspPheSer	440
Db	4677	GGCGCACTTCGACTCGGTGATTAACATGCAATAAGTGTCACCCAGCAAGTCAATTTCAAGC	4736
QY	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	4737	CTTGAACCTTAACCTTCAACATTAGACAAATACAGCTTCCCAAGATGCTGTCTTCCGCACT	4796
QY	461	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
Db	4797	CAACGTCGGGAGAGACTGGACGGGGGAAGCCAGGAGATCTACAGATTTGTGCAACGGGG	4855
QY	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500

```

Db      4857 GAGGGCCCCCTCGGCGATGTTGCACTCGTCGGTCTCTGTGAGTCTATTGACGCGAGGCGT 4916
QY      501 AlaTpyTgTleuThrPProAlaGluThrThrValArgLeuAlaAlaTyrMetAsnThr 520
Db      4917 GCTTGATATGAGCTCAGCGCCCGCGAAGACTTACAGTTAGGCTTACAGAGGTCATGAAACAC 4976
QY      521 ProGlyLeuPProValCysGlnAspHisLeuGluPheThrGluGluValPheThrGlyLeu 540
Db      4977 CCGGAGCGCTCCCGGTGTCAGAGACCATCTTGAAATTTTGGAGGGCGCTTTTACAGGGCTTC 5036
QY      541 ThrHisIleAspAlaHisPheLeuSerGlnThrTyrGlnSerGlyGluAsnLeuProTyr 560
Db      5037 ACTCATATAGATGGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5096
QY      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPheSerTyrAsp 580
Db      5097 CTGGTAGGATACCAAGCCACCGGTGCGCTAAGGCTCAAGGCCCTCCGCCATCGTGGGAC 5156
QY      581 GlnMetTpyTyrCysLeuIleArgLeuLeuPProThrLeuHisGlyProThrProLeuLeu 600
Db      5157 CAGATGTGGAAAGTGTGATTCGGCTCAAGCCACCTCCATGGGCGCAACCCCTCTCTA 5216
QY      601 TyrArgLeuGlyAlaValAlaGlnAspGluValThrLeuThrHisPProValThrTyrIle 620
Db      5217 TACGAGCTGGGCGCTGTTCAAGATGAATCACTGACGACCCAGTCAACCAATATCATC 5276
QY      621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTyrValLeuValGlyGly 640
Db      5277 ATGACATGCAATGTCGGCGGACCTGAGAGTGTCGTCACGACACTGGGTGCTGTTGGCGGC 5336
QY      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5337 GTCTGTGCTGTGTTGGCGCGGTATTTGCTGTCAACAGGCTGCGTGCATATAGTGGCAGG 5396
QY      661 IleValLeuSerGlyLysPProAlaIleIlePProAspArgGluValLeuTyrArgGluPhe 680
Db      5397 GTGCTGTGTGTCGGGAAGCGGCAATCATCTGACAGGAAGTCTTACCGAGAGTTC 5456
QY      681 AspGluMetGluGluCys 686
Db      5457 GATGAGATGGAAGAGTGC 5474

RESULT 6
US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Cho, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
;

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 342..9374
FEATURE:
NAME/KEY: misc_feature
LOCATION: 366
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Arg."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-549-9

```

```

Alignment Scores:
Pred. No.: 0
Score: 3574.00
Percent Similarity: 99.42%
Best Local Similarity: 97.96%
Query Match: 98.78%
DB: 4
Gaps: 0
Length: 9401
Matches: 672
Conservative: 10
Mismatch: 4
Indels: 0

```

US-09-930-591-2 (1-686) x US-08-440-549-9 (1-9401)

Oy 1 MetLaprllePrlrAlrAlrAlaGlnGlnThrArgSlyLeuenglyCyslleleThr 20
:::|||||

Dh 3417 CTGGCCGCAATCACGGCGTACGCCAGACAGAAAGGGCCCTCTAGGGTCATATATACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGlnGlyGlnGlnIleValSerThrAla 40
Dh 3477 AGCTTAATGGCCGGGAGCAAAACCAAGTGAAGGTGAGTCCAGATTGTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleHisGlyAla 60
Dh 3537 GCCCAACCTTCCCTGGCAACCTGTCATATGGGGTGTGCTGACCTGTCTACACAGGGGCC 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetIleThrAsnValAsp 80
Dh 3597 GGAACGAGACATCGCTGACCCCAAGGGTCTGTCTATCCAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Dh 3657 CAAGACCTGTGGGCTGGCCGCTCCGAGAGTACCGCTCATTTGACACCTGACCTTGC 3716
Qy 101 GlySerSerAspLeuIleValThrArgHisAlaAspValIleProValArgArgArg 120
Dh 3717 GGGCTCTCGGACCTTTACTGTGTACAGAGGACCGCATGTCTCCGTCGCGCGCGG 3776
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerIleLeuLysGlySerSer 140
Dh 3777 GGTGATGACAGGGGACGCTGTGTGCCCCGCGCATTTCTTAAAGGAGCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Dh 3837 GGGGGTCCGCTGTTGTGCCCCGCGGACCGCTGGGCATTTTAGGGCCGCGTGTGC 3896
Qy 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGlnThrThrMet 180
Dh 3897 ACCCGTGAAGTGGCTAAGCGGTGAGCTTTATCCCTGAGAACTAGAGAACACATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerIleGlnVal 200
Dh 3957 AGGTCCCGGATGTCAGAGTAACTCTCTCCACAGATAGTCCCGACCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrIleValProAlaAlaValAla 220
Dh 4017 GCTCACTCTCATGTCCCAACAGGCGGAGGCAAAAGACCAAGGTCCCGGCTGACATATCA 4076
Qy 221 AlaGlnGlyIleIleValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Dh 4077 GCTCAGGGCTATAAAGGTGATGACTCAACCCCTGTGTGTGCAACATGGGCTTGTGT 4136
Qy 241 AlaIleThrMetSerIleValHisGlyIleAspProAsnIleArgThrGlyValAlaGlnIle 260
Dh 4137 GCTTACATGTCCAAAGCTCATGGATGATCTTAACATCAGAACCGGGGTGAGAACATTT 4196
Qy 261 ThrThrGlySerProIleThrIleThrIleThrGlyLysPheLeuAlaAspGlyGlyCys 280
Dh 4197 ACCACTGGAGCCCATCACTACTCTCACTACGGCAAGTTCCTTCCGACGGGAGGTGC 4256
Qy 281 SerGlyValAlaIleValAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
Dh 4257 TCGGGGGGCGCTATATGACATATATTTGTGACAGATGCCACTCAGATGCCAATCC 4316
Qy 301 IleLeuGlyIleIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaGlnThrVal 320
Dh 4317 ATCTTGGGCAATCGGCACTGTCTTGTACCAAGAGAGCTCGGGGGGAGAACATGGTGTGTG 4376
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGlnGlnVal 340
Dh 4377 CTGGCCACCGCCACCTCCGGGCTCCGTACTGTGCTGCTCCATCCCAATCGAGAGGTT 4436
Qy 341 AlaLeuSerThrThrGlyGlnIleProPheTrpGlyValAlaIleProLeuGlnAlaIle 360
Dh 4437 GCTCTGCGCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATAC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerIleValLysCysAspGlyLeuAlaAla 380
Dh 4497 AAGGGGGGAGACATCTCATCTTCTGTCAAAAGAAAGTGGCAGAACCTCGCGCA 4556

Qy 381 LysLeuValAlaLeuGlyValAlaAlaValAlaIleTrpArgGlyLeuAspValSerVal 400
Dh 4557 AAGTGTGCTCATTTGGGCATCAATGCGGTGCTACTACCGCGTCTTGAAGTGTCCGTC 4616
Qy 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Dh 4617 ATCCCGACCGCGCGCATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4676
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrGlnThrValAspPheSer 440
Dh 4677 GGGACTTGCATCTCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4736
Qy 441 LeuAspProThrPheThrIleGlnThrIleThrLeuProGlnAspAlaValSerArgThr 460
Dh 4737 CTGACCTCATCTTACCATTTAGACATGACATGACATGACATGACATGACATGACATGACAT 4796
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTrpArgPheValAlaProGly 480
Dh 4797 CAACGTGCGGGCAGACCTGCGCAGGGGGAGCCAGGACATCTACAGATTGTGGACACGGGG 4856
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGlnCysTrpArgAlaGlyCys 500
Dh 4857 GAGGCGCCCTCCGACATGTGTGACTGCTCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 4916
Qy 501 AlaTrpTrpGlyLeuThrProAlaGlnThrThrValArgLeuArgAlaTrpMetAsnThr 520
Dh 4917 GCTTGGTATAGTCTACAGCCCGCGAGACTACAGTTAGCTACAGAGGTACAGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheTrpGlnGlyValPheTrpGlyLeu 540
Dh 4977 CCGGGGCTTCCCGTGTGCGCAGACCATTTGAAATTTGGAGGGCGCTTTTACAGGCGCTC 5036
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrGlySerGlnGlnValLeuAsnProTrp 560
Dh 5037 ACTCATATATATATCCCATCTTATCCCAAGACAAAGCAGAGTGGAGAACTTCTCTTAC 5096
Qy 561 LeuValAlaIleTrpGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Dh 5097 CTGGTAGCGTACCAACCGACCGGTGTGCTAGAGGCTCAAGCCCTCCCATGTGTGGAC 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Dh 5157 CAGATGTGAAGTGTGATTCGCTCAAGCCACCTCCATGTGGCCAAACCCCTGTCTA 5216
Qy 601 TyrArgLeuGlyValAlaGlnAsnGlnValThrLeuThrHisProValThrIleGlyIle 620
Dh 5217 TACAGACTGGCGCTGTTCAGAAATGAAATCACCTTACAGCACCCAGTCAACAAATGATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGlnValIleThrSerThrTrpValLeuValGlyGly 640
Dh 5277 ATGACATGATGTGCGGCCGACCTCGAGGTGTGTACAGACACCTGGGTGTCTGTGGCGGC 5336
Qy 641 ValLeuAlaAlaLeuAlaIleValIleValIleValIleValIleValIleValIleVal 660
Dh 5337 GTCCTGAGCTGTTGGCCGCGGTATGCTGTCAACAGGCTGCGTGTGATAGTGGGAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlnValLeuTrpArgGlnPhe 680
Dh 5397 GTGCTTGTTCGGGAAAGCCGCAATCATACCTGACAGGAGAAATCTTACACGAGATTC 5456
Qy 681 AspGluMetGlnGlnCys 686
Dh 5457 GATGAGTGGAGAGTGC 5474

RESULT 7
US-08-823-895A-25
; Sequence 25, Application us/08823895A
; Patent No. 643159
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson
; TITLE OF INVENTION: Compositions And Methods For
; TITL OF INVENTION: Treatment Of Hepatitis C Virus-Associated Diseases

NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 66 E. Main Street
 CITY: Marlton
 STATE: NJ
 COUNTRY: USA
 ZIP: 08053
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 COMPUTER: IBM 486
 OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/823,895A
 FILING DATE: March 17, 1997
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/453,085
 FILING DATE: May 30, 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/945,289
 FILING DATE: September 10, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: ISPH-0203
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 810-1454
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9401
 TYPE: Nucleic
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: NO
 US-08-823-895A-25

 Alignment Scores:
 Pred. No.: 0 Length: 9401
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 Gaps: 0
 DB: 4
 US-09-930-591-2 (1-686) x US-08-823-895A-25 (1-9401)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 3417 CTGGCGCCCATCAACGGCTGACGCCGACGACGACGAGGGGCTCTTAAAGGTGATATACACC 3476
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
 DB 3477 AGCTTAACGTGGCCGGGCAAAAACCAAGTGGAGGTGAGTCCAGATTGTGTCAACTGCT 3536
 QY 41 AlaGlnThrPheLeuAlaIleThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
 DB 3537 GCCCAACCTTCCTGGCAACGTCATGAGGTGTCTGACTGTCTTACCAAGGGCC 3596
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 3597 GGAAGAGAGACATCGCGTCAACCCAAAGGTCTCTGTCAATGATATACCAATGTAGAC 3656
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 3657 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTACCCGCTCATTAACACCCCTGCACTTGC 3716
 QY 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 3717 GGCTCCTCGGACCTTACTGAGTCAAGAGCACGCCGATGTCTCCGTCGGCGCGGCGG 3776

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
 DB 3777 GGTGATAGAGAGGAGCGAGCTGTGTGCGCCGCGCCATTTCTCTTAAGAGGCTCTTCG 3836
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3837 GGGGGTCCGCTGTGTGCCCCGGGGGACCGCTGGCGATATTAGAGCCGGGTGTGCG 3896
 QY 161 ThrArgGlyValAlaIleAspAlaValAspPheIleProValGluSerLeuGlnThrMet 180
 DB 3897 ACCCTGAGTGGCTTAAGCGCGTGAATTTATCCCTGTGGAGAACCTTAGACCAACCATG 3956
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 3957 AGGTCCCGCGTGTCAAGGATTAATCTCTCCACAGATAGTCCCAAGCTTCCAGAGTTCG 4016
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIleTyrAla 220
 DB 4017 GCTCACTTCATGCTCCACAGGACGGGCAAAAGACCAAGGTCCCGGCTGCATATGCA 4076
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly 240
 DB 4077 GCTCAGGCTATTAAGTGTCTAATCAACCCCTGTGTGCAACACTGGGCTTTGGT 4136
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 4137 GCTTACATGTCCAAAGGCTCATGAGATCGATCTTAACATGAGACCGGGGTGAACAAATT 4196
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 4197 ACCACTGGACGCCCATCAACGTAATCTCACTCAAGCAATTTCTTCCGACGGCGGTGC 4256
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 4257 TCGGGGGGCGCTTATGACATAATTAATTGTGACGAGTGCACCTCCACGATGCCACATCC 4316
 QY 301 IleLeuGlyTyrIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 4317 ATCTTGGGCAATCGGCACTGTCTTGAACCAAGACATGCGGGGCGAGACTGTGTGTCG 4376
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGlnVal 340
 DB 4377 CTGGCCACCGCACCCCTCGGGCTCGTCACTGATGCCCATCCCAATCGAGGAGTTC 4436
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGlnAlaIle 360
 DB 4437 GCTCTGTCACACACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCGAAGTAAATC 4496
 QY 361 LysGlyGlyArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 DB 4497 AAGGGGGGAGACATCTCATCTTGTGATTCAAAGAGAGTGCACAGCACTCCGCGCA 4556
 QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 4557 AAGCTGTGCTGATGGGCAATGACGCGGTCCCTTACACCGCGGTCTTGAAGTGTGCGTC 4616
 QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
 DB 4617 ATCCGACACAGCGCGCATGTTGTGTGTGTCGCAACGATCCCTCATGACCGGCTTAATCC 4676
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 4677 GGCACCTTCACTCGGATGATGACGATCAATGCTGTGTCAACGACGATGATTTTCAGC 4736
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 4737 CTTAACCTTCACTTCAATTTAGACAAATACCGTCCCGAGATGTGTCTCCCGGCACT 4796
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 4797 CAACGTGCGGGCAGAGATGCGACAGGGGGAACCGAGCATTCACGATTGTGGCCACGGGG 4856
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500


```

Db      1863 GCTCAGGGCTAATAAGTGTGTAATCTCAACCCCTCTGTGCTGCAACACTGGGCTTGTG 1922
Qy      241 AATyTMeSerlysalahisgilyleaspproasnilleargthrglyValargthrlle 260
Db      1923 GCTTACATGTCACAGGCTCATGGGATGATCTTAACATCAGAGACGGGGGAGAACAAAT 1982
Qy      261 ThThrglySerproiliethtyrSerThThrglylyspheleuilaaspgilylyCys 280
Db      1983 ACCACTGGCAGCCCATCAGTACTCTCACTCAGCAAGTTCCTTGCCAGCGGGGGTGC 2042
Qy      281 SercllygilyaIatYraapillelleCysAspglyCysHisSerThraapalathSer 300
Db      2043 TCGGGGGGGCTTATGACATTAATTTGTGACAGACTGCCACTCAAGATGCCACTCC 2102
Qy      301 IleleugilylleglyThraValleuaspGlnalagluThralagilylaargleuthVal 320
Db      2103 ATCTGGGCACTGGCAGCTGCTTGACCAAGCAAGACTGGGGGGGAGAACTGGTGTG 2162
Qy      321 LeuAlaThraAlaThrrpropGlySerValThraValProHisproAsnillegluGluVal 340
Db      2163 CTGCGCACCGCCACCCCTCGGGCTCGGTCACTGTGCCATCCCAATCCAGAGAGGTT 2222
Qy      341 AlaleuSerThThrglygluIleproPheTyrGlylysalalleproleugluAlalle 360
Db      2223 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTATC 2282
Qy      361 LysglylyarghisleuilepheCysHisSerlyslslyCysAspglyleuAlaAla 380
Db      2283 AAGGGGGGAGACATCTCATCTCTGTCTGATTCAAAGAGAGTCCGAGAACTGGCGCA 2342
Qy      381 LysleuValAlaleuGlyValAsnAlaValAlaTyrTyrArgGlyleuAspValSerVal 400
Db      2343 AAGCTGTGATGGAGCATCAATGCCGTGGCTTACCCGCTTGAAGTCTCCGTC 2402
Qy      401 IleproThrSerGlyAspValValValValAlaThraAlaaleuMetThrglyPheThr 420
Db      2403 ATCCGACCAACGGGGATGTTGTGTGTCGCAACCGATGCCCTCATGACCGGCTAATACC 2462
Qy      421 GlyaspPheaspservalilleaspcysasnhrCysValThThrglyValaspPheSer 440
Db      2463 GCGGACTTCGACTCGGTATGACATGCAATAGCTGTGCACCCAGACAGTGAATTCAGC 2522
Qy      441 LeuaspProThrrPheThrllegluThrlleThrlleuProGlnAspAlaValSerArgThr 460
Db      2523 CTTGACCTTACCTTACCATTCATGAGCATCAOGCTCCCCAGAGATGCTGTCCCGGACT 2582
Qy      461 GlnargarglyargThrglyargGlylyspProGlyllelyrArgPheValAlaProGly 480
Db      2583 CAACGTCGGGGGAGAGCTGGCAGGGGAGAACGACGATCTCAAGATTTGTGGCAACGGGG 2642
Qy      481 GlnargproserGlyMetPheaspservalilleuGlyCysGlyTyrAspAlaglyCys 500
Db      2643 GAGCGCCCTCGGAGATGTCAGCTGTCGTCCTCTGTGATGCTTAAGCAGGCTGT 2702
Qy      501 AlaTyrGlyleuThrrProAlagluThrrValargleuArgAlaTyrMetAsnThr 520
Db      2703 GCTTGATGAGCTACACGCGCCGAGACTACAGTTAGGCTTCAGCGCAATGAACACC 2762
Qy      521 ProGlyleuProValCysGlnAspHisleuGluPheTyrGluGlyValPheThrglyleu 540
Db      2763 CCGGGGGCTTCGGGTGGCAGAGCATCTTGAATTTGGAGGGGCTCTTAAACAGGCTTC 2822
Qy      541 ThThrsilleaspAlaHisPheleuSerGlnThrllysglnSerGlyGluAsnleuProTyr 560
Db      2823 ACTCAATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACTTCTTAC 2882
Qy      561 LeuValAlaTyrGlnAlaThrrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db      2883 CTGGTAGCTACCAAGCCACCGGTGGCTAGGGGCTCAAGCCCTCCCATCGTGGGAC 2942
Qy      581 GlnMetTyrlyCysleuileargleuIleargleuIleProThrlleuHisGlyProThrrProleu 600
Db      2943 CAGATGTGAGAGTGTGTGATTCGCTCAAGCCCACTTCATGGGGCAACACCCCTGTCTA 3002

```

```

Qy      601 TyrArgleuGlyAlaValGlnAsnGluValThrlleuThrrHisProValThrllyTyrIle 620
Db      3003 TACAGACTGGCGCTCTTATGAATGAATACACCTGACGACCAAGTACCAAAATATATC 3062
Qy      621 MetThrCysMetSerAlaAspleuGluValValThrrSerThThrrValleuValGlyly 640
Db      3063 ATGACATGATGTGGCGCCGACCTGGAGGTGTCACAGACCACTGGGTGCTGTTGGCGGC 3122
Qy      641 ValleuAlaAlaLeuAlaAlaTyrCysleuSerThrglyCysValAlalleValGlyArg 660
Db      3123 GTCTGGCTGCTTTGGCGCGATTTGCTGTCAACAGGCTGCTGCTCATAGTGGCAGG 3182
Qy      661 IleValleuSercllylyspProAlallelleproAspArgGluValleuTyrargGluPhe 680
Db      3183 GTGCTTGTTCGGGAGAGCGGCAATCATCTGACAGAGGAAGTCTTACCGAGATTC 3242
Qy      681 AspGluMetGluGluCys 686
Db      3243 GATGAGATGAGAGAGTGC 3260

```

RESULT 9

```

US-08-444-818-88
; Sequence 88, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
; US-08-444-818-88

```

Alignment Scores:

```

Pred. No.: 0
Score: 3570.00
Percent Similarity: 99.27%
Best Local Similarity: 97.81%

```

Length: 8316

Matches: 671

Conservative: 10

Mismatches: 5

Query Match: 98.67% Indels: 0
DB: 3 Gaps: 0
US-09-930-591-2 (1-686) X US-08-444-818-88 (1-8316)
QY 1 MetAlaProIleThrAlaIAlaGlnGlnThrArgIlyLeuLeuGlyCysIleIleThr 20
DB 2734 CTGGGCGCCATCAACGGGTACGGCCAGCAGCAAGGGGCTCTAGGGTCCATATCAAC 2793
QY 21 SerLeuThrGlyArgAspIysAsnGlnValGluGlyValGlnIleValIleSerThrAla 40
DB 2794 AGCTTAACCTGGCCGGGCAAAAACCAAGTAGAGGTAGGTCCAGATTGTGTCACTGCT 2853
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIlyrhisGlyAla 60
DB 2854 GCCCAACCTTCTGGGCAACGTGCATCAATGGGGTGTGCTGACTGTCTACCAAGGGGCC 2913
QY 61 GlyThrArgThrIleAlaSerProIyGlyProValIleGlnMetIyThrAsnValAsp 80
DB 2914 GGAAACGAGACATCGGTCACCCCAAGGTCCTGTCTCCAGATGTATACCAATGTAGAC 2973
QY 81 GlnAspLeuValGlyTrpProIaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 2974 CAAGACCTTGTGGGCTGGCCGCTCGCAAGTAGCCGCTCATTTGACACCTGCACTTGC 3033
QY 101 GlySerSerAspLeuIyLeuValThrArgHisAlaAspValIleProValIleArgArg 120
DB 3034 GGCTCTCGGACCTTTAATCTGTGTACGAGGACGCGCATGTCAATCCCGTGGCCGGCGG 3093
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerIyLeuIyGlySerSer 140
DB 3094 GGTAATAGCAGGGGAGCCTGTGTGCTGCCCCGGCCCTTTCTAATTGAAGGCTCTCG 3153
QY 141 GlyGlyProLeuLeuCysProIaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3154 GGGAGGTCTGCTGTGTGCCCCGGGGGACGCGTGCGCATATTAGGGCGCGGTGTGC 3213
QY 161 ThrArgGlyValAlaIyAspValIleAspPheIleProValGlySerLeuGlnThrThrMet 180
DB 3214 ACCGTGTAGCTGTAGGCGGTGAGCTTATTCCTGTGAGAACTTAGAGACACATG 3273
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerIyGlnVal 200
DB 3274 AGGTCCCGGTGTGTACGAGTAACTCTCTCCACAGTAGTGGCCCGACCTTCCAGGTG 3333
QY 201 AlaHisLeuHisAlaProThrArgIySerGlyIySerThrIyValProIaAlaIyValAla 220
DB 3334 GCTCACTCCATGCTCCACAGGAGCGGCAAAAGCACCAAGGTCCCGGCTGCAATAGCA 3393
QY 221 AlaGlnGlyIyIyValIyValIleValIleAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 3394 GCTCAAGGCTATAGGTAGTACTCAACCCCTCTGTCTGCTCAACACGCGCTTGTGT 3453
QY 241 AlaIyMetSerIyAlaHisGlyIleAspProAsnIleArgThrGlyValIleArgThrIle 260
DB 3454 GCTTAACATGTCCAAAGGCTCATGGATGATCTTAACATCAAGACCGGGGAGAAACAATT 3513
QY 261 ThrThrGlySerProIleThrIyTrsSerThrIyGlyIySphLeuAlaAspGlyIyCys 280
DB 3514 ACCACTGGAGCCCAACACAGTACTCCACTACGGCAAGTTCCTTGGCAGCGGGGTGC 3573
QY 281 SerGlyValIyAlaIyAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 3574 TCGGGGGGCGGTTATGACATATATTTTGTGACAGTGCACCTCCACGAGATGCCACATCC 3633
QY 301 IleLeuGlyIleGlyThrValIleuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 3634 ATCTGGGCAATCGGCACTGTCTTGACCAAGACAGACTCGGGGGGAGACATGTTGTG 3693
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 3694 CTGGCAACCGCCACCTCGGGCTCGGTCACTGTGCCCATCCCAACATCGAGAGGTT 3753

QY 341 AlaLeuSerThrThrGlyGluIleProPheIyIyValAlaIleProLeuGluAlaIle 360
DB 3754 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGGAGAGGTATCCCTCGAAGTAAATC 3813
QY 361 IyGlyGlyArgHisLeuIlePheCysHisSerIyValIyCysAspGlyLeuAlaIle 380
DB 3814 AAGGGGGGAGACATTCATCTTGTGTCAATCAAAAGAGGACGACAACTCGCGGCA 3873
QY 381 IyLeuValAlaIleuGlyValAsnAlaValAlaIyIyTrsArgGlyLeuAspValSerVal 400
DB 3874 AAGCTGTGCACTGGGATGATCAATGCGGTGCTTACCGGGCTTACGTGTGCGTGC 3933
QY 401 IleProThrSerGlyAspValValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
DB 3934 ATCCCAACAGCGGCGATGTGTGCTGTGCGCAACGAGTCCCTCATGACCGGCTATACC 3993
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 3994 GGGACTTGACTCGGTGATGATGATGCAATACGTGTGTACCCAGACAGTCAATTCAGC 4053
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4054 CTTGACCTTAACCTTCAACATTAAGACATCAACGCTCCCAAGATGCTGTCTCCCACT 4113
QY 461 GlnArgArgGlyArgThrGlyArgIyIySerProGlyIleIyTrsArgPheValAlaProGly 480
DB 4114 CAACGTGGGGGAGGACTGGAGGGGGAAGCAGGAGCATCAAGATTTGTGGCACCGGGG 4173
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysIyTrsAspAlaGlyCys 500
DB 4174 GAGCGCCCTCGGCGATGTGTCACTGTGCTGCTGTGAGGTATAGACGAGGCTGT 4233
QY 501 AlaTrpIyGlyIleuThrProIaGluThrThrValIleArgLeuArgAlaIyThrMetAsnThr 520
DB 4234 GCTTGATAGACTCAACGCCCGCCAGACTTAAGTTAGGCTAGAGCGTATAGAACACC 4293
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
DB 4294 CCGGGGCTTCCGCTGTGCCAGGACCATTTGAATTTTGGAGGGGCTTTTACAGGCTTC 4353
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIySglnSerGlyIyAsnLeuProIy 560
DB 4354 ACTCATATAGATGCCCACTTTTATTCACAGCAAAAGCAGAGTGGGAGAACCTTCTTAC 4413
QY 561 LeuValAlaIyTrsGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 4414 CTGGTAGCTTACCAAGCCACCGTGTGCTTAAAGGCTCAAGCCCTCCCATGTGGAGC 4473
QY 581 GlnMetTrpIyCysLeuIleArgLeuIySProThrLeuHisGlyProThrProLeuLeu 600
DB 4474 CAGATGTGAAGTGTGATTGATTCGCTCAAGCCCACTTCATGGGCAACCCCTGTGTA 4533
QY 601 TyrArgLeuGlyValAlaGlnAsnGlyValIleThrLeuThrHisProValThrIyTrsIle 620
DB 4534 TACAGACTGGGCGCTGTGATGAATGAATCACCTTGAACCAACCACTCAACAAATACATC 4593
QY 621 MetThrCysMetSerAlaAspLeuGluValIleThrSerThrTrpValLeuValGlyIy 640
DB 4594 ATGACATGATGTGGCCCACTGAGGTGTGTACAGCAACCTGGGTGTCTGTGGCGGC 4653
QY 641 ValLeuAlaAlaLeuAlaIleAlaIyTrsCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 4654 GTCTGTGCTGTGTGGCCGCTATGCTGTCAACAGGCTGGGTGTCTAATGTGGCAAG 4713
QY 661 IleValLeuSerGlyIySProAlaIleIleProAspArgGluValLeuIyTrsArgIyPhe 680
DB 4714 GTCTGTGTTCGGGGAAGCCGCAATCATATCTGACAGGGAAGTCTTCAACGAGATTC 4773
QY 681 AspGluMetGluGluCys 686
DB 4774 GATGAGATGAAAGATGCT 4791

RESULT 10

US-08-444-818-122
Sequence 122, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Ruter, William J.
TITLE OF INVENTION: NANBY Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hardin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 9185 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-444-818-122

Alignment Scores:
Pred. No.: 0 Length: 9185
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.67% Indels: 0
Gaps: 3

US-09-930-591-2 (1-686) x US-08-444-818-122 (1-9185)

QY 1 MetAlaprollethralaTyAlaGlnGlnThraArgglyLeuLeuglyCysIlelleThr 20
Db :::::
3395 CTGGGGCCCATCAAGCGCTACGCCACAGAGGGGCTCTCAAGGTGATATACCC 3454
QY 21 SerleuthrArglyAspIleAsnGlnValGluGlyGluValGlnIleValSerThra 40
Db 3455 AGCCTAAGTGGCCGGGCAAAAACCAAGTGAAGGTGAGTCCAGATTGTGTCACCTGCT 3514
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyRHisGlyAla 60
Db 3515 GCCCAAACTTCTCTGGCAAGTGCATCATGGGGTGTCTGACGTCTACACAGGGGCC 3574
QY 61 GlyThraArgThrIleAlaSerProIysGlyProValIleGlnMetTyThraAsnValAsp 80
Db 3575 GGAAGAGAGACCATCGCTCAACCAAGGATCTCTGTATCATCGATGTATACCAATGTAGAC 3634
QY 81 GlnAspLeuValAlaGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGACCTTGTGGGCTGGCCGCTCGCAAGGTAGCCCTCATTTAGACCCCTTGACCTTGC 3694

QY 101 GlySerSerAspLeuTyIleuValThrArgHisAlaAspValIleProValArgArg 120
Db 3695 GGCTCTCGGACCTTACTTGCTGCAGGAGCAGCGCATGTGATTCCTCGTGGCGCGG 3754
QY 121 GlyAspGlyArgGlySerleuLeuSerProAlaGProIleSerTyIleuIleGlySerSer 140
Db 3755 GGTGATAGCAGGGGAGCTGCTGTGCGCCCGGCAATTTCTTACTTGAAAGGCTCTCG 3814
QY 141 GlyIleProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGAGGTGCGCTGTGTGCCCCCGGGGACGCGGTGGGATTTTAGGGCCCGGTGTGC 3874
QY 161 ThrArgGlyValAlaAlaValAspPheIleProValGlySerLeuGlnThrThre 180
Db 3875 ACCCGTAGAGTGAAGCGGTGAGCTTATCTCTGTGAGAGACTGAGCAACAGCAG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyGlnVal 200
Db 3935 AGGTCCCGGTGTTCACGATTAATCTCTCCACAGTAGTCCCGAGGCTTCCAGGTG 3994
QY 201 AlaHisLeuHisAlaProThrGlySerGlyIleYsSerThrIleValProAlaAlaTyAla 220
Db 3995 GCTACCTTCAGTCTCCACAGGACGGCAGAAAGCAGCAAGTCCCGGCTGCATATGCA 4054
QY 221 AlaGlnGlyTyIleValIleuValIleuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGCGTATPAGGTGCTAGTACCAACCTCTGTGTGCTGCAACCTGGGCTTTGGT 4114
QY 241 AlaTyMetSerIleValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAAAGGCTCATGGATGATCACTTACATAGAGCCGGGTGAGAACATTT 4174
QY 261 ThrThrGlySerProIleThrTySerThrTyArgIleYsPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 4234
QY 281 SerGlyGlyAlaTyAspIlellelleCysAspGlyCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGGCGGTATGATCAATTAATTTGTGAGAGTGGCCACTCCAGATGCAATGCC 4294
QY 301 IleLeuGlyIleGlyThrValIleuAspGlnAlaGlyThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCACTGGCACTGTCTTGTGACAGAGAGACCTCGGGGGGAGACGTGTG 4354
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluIleVal 340
Db 4355 CTGGCAGCGCCACCCCTCGGGCTCGGTACATGTCGCCCATCCCAATCCGAGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyIleuIleProPheTyArgIleYsAlaIleProLeuGlnAlaIle 360
Db 4415 GCTCTGTCCACCAACCGAGAGATCCCTTTTACGGAAGGCTATCCCTCGAGATATC 4474
QY 361 LysGlyGlyIleArgHisLeuIlePheCysHisSerIleYsCysAspGluLeuAlaIle 380
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTCAATTAAGAAAGAGCGAGCAACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleTyThrArgGlyLeuAspValSerVal 400
Db 4535 AAGGTGTGCAATTTGGGATCAATGCGGTGCTACTACCGCGGTCTTGACGTGTCCGTC 4594
QY 401 IleProThrSerGlyAspValIleValIleValIleThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCGACAGCGCGGATGTGTGTGTGGCAACGATGCCCTCATGACCGGCTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGGACATTCATCTCGGATGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCAACATTTAGACATACCGCTCCCGAGATGTGTCTCCCGCACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyIleYsPheProGlyIleTyThrArgPheValAlaProGly 480

DB 4775 CAACGTCGGGCGAGATGCGAGGGGGAGCCAGGCACTTACAGATTGTGGACCGGG 4834
QY 481 GUAAGPProserGlyMetPheAspSerSerValLeuGlyCysTyrAspAlaGlyCys 500
DB 4835 GAGCGCCCTCCCGGCAATGTTGCACTGCTCCGTCTGTAGTGTCTTACAGAGGCTGT 4894
QY 501 AlaTPTyGlyLeuThrProAlaGlyThrThrValArgLeuAlaGlyMetAspThr 520
DB 4895 GCTTGGATAGAGCTCAAGCCCGGAGACATACAGATTAGCTCAAGAGGTACAGAACCC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGlnPheThrProGlyValPheThrGlyLeu 540
DB 4955 CCGGGGCTTCCCGTGTGCGAGACCACTTGAATTTGGAGGGCGCTTTTACAGGGCTC 5014
QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrArgGlnSerGlyValLeuLeuProTyr 560
DB 5015 ACTCATATAGATGCCCTTCTATCTTATCCACAGAACAGAGTGGGAGAACCTTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProPProSerTyrPasp 580
DB 5075 CTGGTAGCGTACCAAGCCAGCGTGTGCTAGGGCTCAAGCCCTCCCATCGTGGGAC 5134
QY 581 GlnMetTyrGlyCysLeuIleArgLeuLeuProThrLeuHisGlyProThrProLeuLeu 600
DB 5135 CAGATGTGAAGTGTGATTGCTGCTCAGCCCACTCCATGGGCGCAACGCCCTGCTA 5194
QY 601 TyrArgLeuGlyValAlaGlnAspGlyValThrLeuThrHisProValThrLysTyrIle 620
DB 5195 TACAGACTGGGCGCTGTTCAAGATGAAATACCCCTTACGAGCCCACTCAGCAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGlyValThrSerThrTyrValLeuValGlyGly 640
DB 5255 ATGACATGATGTCGCGGCGAGCTGAGGTGCTCAGCAGGACCTGGGGCTCGTTGGGGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 5315 GTCTGTGCTGCTTGGCGCGGTATGCTGTCAACAGGCTGCGGTGCTCATATGAGGGGAG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGlyValLeuTyrArgGlyPhe 680
DB 5375 GTGTCTTGTCCGGAGGCGGCAATCATCTGACAGGAGTCTTACCGAGGTTCC 5434
QY 681 AspGluMetGluGlyCys 686
DB 5435 GATGATGTGAAGATGTC 5452
RESULT 11
US-08-444-818-123/C
Sequence 123, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESSES:
ADDRESSES: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hardin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110,002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 9185 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ANTI-SENSE: YES
US-08-444-818-123
Alignment Scores:
Pred. No.: 0 Length: 9185
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.67% Indels: 0
Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-123 (1-9185)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 5791 CTGGCCCCCATACAGCGCGTACGCGCCGAGACAGAGGGCGCTCTAGGGTGCATTAACACC 5732
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 5731 AGCTTACTGGCCGGGACAAAACAGTGAAGGTGAAGTCCAGATTGTGCACTGCT 5672
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
DB 5671 GCCCAACTTCTCTGGCAACGTCATCATGGGGTGTCTGGAGCTGTCAACACGGGGCC 5612
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 5611 GGAACGAGGACCATCGCTCACCCCAAGGCTCTCATCATCAATGATACATGATGAC 5552
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 5551 CAAGACTTGTGGGCTGGCGCGCTCCGCAAGTAGCGCTCATGACACCTGCACTTGC 5492
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 5491 GGCTCTCGGACCTTACCTGTGTGTCAGAGGACACCGCATGTCACTTCCCGCCGGCGG 5432
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 5431 GGTGATAGAGGGGAGCGCTGTGTGCGCCCGGCAATTTCTTCACTTGAAGGCTCTCGG 5372
QY 141 GlyGlyProLeuLeuLysCysProAlaGlyHisAlaValGlyTyrPheArgAlaAlaValCys 160
DB 5371 GGGGCTCGCTGTGTGCGCCCGGGGACGGCGGTGGGATTTTGAAGGCGCGGGTGC 5312
QY 161 ThrArgGlyValAlaLysValAlaValAspPheIleProValGlnSerLeuGlnThrThrMet 180
DB 5311 ACCCGTAGAGTGGCTAGAGCGGTGACCTTATCCCTGTGAGAACTTAGACACACCATG 5252
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 5251 AGTCCCGCGTGTTCACGGATGATCTCTCCACCAAGTAGTGTCCCGCAGAGCTTCCAGGTG 5192
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrThrValProAlaTyrAla 220
DB 5191 GCTCACTTCATGCTCCCAAGGCGGCAAAAGCAACCAAGATCCCGGCTGCATATGCA 5132

QY 221 AlaGInGlyTyrIleValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 5131 GCTCAGGGCTAATAGGGCTAGTACTCAACCCCTGTTGCTGCACACCTGGCTTGGT 5072
 QY 241 AlaTyrMetSerIleValIleGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 5071 GCTTACATGTCAGAGGCTCATGGATCGATCTTACATCGACGCCGGGGTGAACAACAT 5012
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyIleCys 280
 DB 5011 ACCACTGGCAGCCCATCACTACTCCACTACGGCAAGTCTCTTGGCGCGCGGGTGC 4952
 QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlnCysHisSerThrAspAlaThrSer 300
 DB 4951 TCGGGGGCGGTTATGACATATATATTGACAGTGCACATCCACGAGAGCCACATCC 4892
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 4891 ATCTTGGGCATCGGCATCTGCTTGCACCAAGAGAGCTCGGGGGGGAAGCTGTTGTG 4832
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 4831 CTGGCCACCGGCACCCCTCCGGGCTCCGTCACGTGCTCCATCCCAATCGAGAGGTT 4772
 QY 341 AlaLeuSerThrThrThrGlyGluIleProPheThrGlyIleValIleProLeuGluAlaIle 360
 DB 4771 GCTCTGTCACACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAATATC 4712
 QY 361 LysGlyIleArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIle 380
 DB 4711 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAGAAGAAGTCCGACGAAGTCCGCGCA 4652
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 4651 AAGCTGGTCGATGGGCATCAATGCGTGCTTACACCGCGGCTTGAAGTGTCCGTC 4592
 QY 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 4591 ATCCCGACAGCGGCGATGTTGTGTCGTGCGCAACCATGCTCCATGACCGGCTATACC 4532
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 4531 GGGGACTTCGACTCGGTGATGACTGCAATACGTGTGTCACCCAGACAGTGCATTCAGC 4472
 QY 441 LeuAspProThrPheThrIleGluThrIlePheThrLeuProGlnAspAlaValSerArgThr 460
 DB 4471 CTTCGACCTTACCTTCACCATGAGACATCAAGTCCGCCAGAGTGTGTCCGCGACT 4412
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 4411 CACGCTCGGGGCGAGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG 4352
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 4351 GAGCGCCCTCCGCGCATGTCGACTGTCGCTCCCTGTGATGATGAGCGAGCGCGT 4292
 QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuAspAlaTyrMetAsnThr 520
 DB 4291 GCTTGGTATGAGCTACCGCCGCGAGACTACAGTTAGGCTACGAGCTTACATGACACC 4232
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 DB 4231 CCGGGGCTTCCCGTGTGCGAGACATCTTGAATTTTGGAGGGGCTCTTACAGGCTTC 4172
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 4171 ACTCATATGATGCCACACTTCTATCCAGCAAAAGAGAGTGGGAGAAACCTTCTTAC 4112
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
 DB 4111 CTGGTAGCGTACCAAGCCACCGGTGCGCTTAGGGCTCAAGCCCTCCCTCCATCGTGGAC 4052
 QY 581 GlnMetTyrPyrCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600

DB 4051 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCCAACCTCCATAGGGCCCAACACCCCTGCTA 3992
 QY 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrIleTyrIle 620
 DB 3991 TACAGATCGGGGCGGTTCAGATGAATATCAACCTGAGCACCCAGTCAACCAATACATC 3932
 QY 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrTyrValLeuValGlyIle 640
 DB 3931 ATGACATGATGATTCGGCCGACCTGAGAGGTGTCACAGACACTGGGTGCTCGTGGCGGC 3872
 QY 641 ValLeuAlaIleLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 3871 GTCCTGCTGCTTGGCGCGGTATGCTGTTCACAGCGCTGCGTGCATATGTGGCAGG 3812
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 3811 GTGCTTGTTCGGGAAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAAGATTCC 3752
 QY 681 AspGluMetGluGluCys 686
 DB 3751 GATGATGGAAGAGTGC 3734
 RESULT 12
 US-09-881-239-2
 ; Sequence 2, Application US/09881239
 ; Patent No. 6630298
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIBEN, David Y.
 ; APPLICANT: ARCANDEL, Phillip
 ; APPLICANT: TANDESKIE, Laura
 ; APPLICANT: GEORGE-NASCIMENTO, Carlos
 ; APPLICANT: COIT, Doris
 ; APPLICANT: MEDINA-SELBY, Angelica
 ; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
 ; FILE REFERENCE: 2302-16073 / P16073-003
 ; CURRENT APPLICATION NUMBER: US/09/881,239
 ; CURRENT FILING DATE: 2001-06-14
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 2058
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: representative NS3/4a conformational antigen
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2058)
 US-09-881-239-2
 Alignment Scores:
 Pred. No.: 0 Length: 2058
 Score: 3565.00 Matches: 671
 Percent Similarity: 99.13% Conservative: 9
 Best Local Similarity: 97.81% Mismatches: 6
 Query Match: 98.54% Indels: 0
 DB: 4 Gaps: 0
 US-09-930-591-2 (1-686) x US-09-881-239-2 (1-2058)
 QY 1 MetaLabProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCCCATACAGGGGTAGCGCCACAGCAAGGGGCTCTCTAGGGGTCAATATATACC 60
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGlyGluValGlnIleValSerThrAla 40
 DB 61 AGCTTAATGCGCGGACAAACCAACAGTGAAGGTAGGTCCAGATTGTGTAACAGTCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyVal 60
 DB 121 GCCAAACCTTCCTGGCAACGTGATCAATGGGGGTGTGTGACTGTCTACACAGGGGCC 180

QY 61 GlyThrArgThrIleAlaSerProValIleGlnMetTyrThrAsnValAsp 80
DB 181 GGAACGAGGACCAATCGGTCACCCAGAGTCTGTCACAGATGATACCAATGTAGAC 240
QY 81 GluAspLeuValGlyTTPProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAAGACTTGTGGCTGGCCCGCTCCGCAAGTACGCAATGATTAACACCTGACTGTC 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCCCTCGGACCTTTACTGTCACAGAGGACCGCATTCATTCGCGCCGCGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlySerSer 140
DB 361 GGTGATGCGAGGGGCGACCTGCTGCTCCCGCCCATTTCTTCAAGAAAGGCTCTCCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheAspGlyAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGCCCCGCGGGCAGCGCGTGGCATTTAGGGCCGCGGTGTC 480
QY 161 ThrArgGlyValAlaValAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCCGTGAATGGCTAAGCGGTGAGTCTTATCCCTGTGAGAACTAGAGACACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGGTGTTCCAGATTAATCTCTTCCACAGTAGTCCCGCCAGACCTTCCAGGTG 600
QY 201 AlaHisIleuHisAlaProThrArgIleSerGlyIleSerThrIleValProAlaAlaTyrAla 220
DB 601 GCTCACTCCATGCTCCACAGGAGCGGCAAAAGCCAAAGTCCCGCTGCAATAGCA 660
QY 221 AlaGlnGlyTyrIleValIleValIleuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTAATAGTGGTACGTAACCTCCCTCTCTGTCGACACCTGGCTTGTGT 720
QY 241 AlaTyrMetSerIleValHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTGCCAAGGCTCATGGGATGATCTTAACATCAGACCGGGGGAGAAACAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyGlyCys 280
DB 781 ACCACGTGGACCCCAACACGTAACCTCCACGCAAGTCTCTTCCGACGCGGGTGC 840
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGlyCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGGCGCTTATGACATTAATTTGTGACAGAGGCCATCCACGAGATGCCATCC 900
QY 301 IleLeuGlyIleGlyThrValIleuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCAATGGCACTGTCTTGAACAAGACACTCGGGGGGAGAGATGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTGCGCACCGCACCCCTCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGAGGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyIleValIlePheLeuGluAlaIle 360
DB 1021 GCTCTGCCACCAACCGAGAGATCCCTTTTACGGCAGGCTTATCCCTCGAAGTATC 1080
QY 361 LysGlyGlyValArgHisLeuIlePheCysHisSerIleValIleCysAspGluLeuAlaIle 380
DB 1081 AAGGGGGGAGACATCTCATCTTCTGTCAATTAAGAAAGAGGCGAGCAATCCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAlaAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGCTCCATTTGGCATCAATGCGGTGGCTTACTACCGCGCTTGAACGTCCTGC 1200
QY 401 IleProThrSerGlyAspValIleValIleValIleThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGCCCATCGGCATGTTGTGCTGTGTGCAACGATGCCCTCATGACCGGCTATAC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

DB 1261 GGGACTTGACCTCGGTGATGACTCAATACGTGTGTCACCCAGACAGTCAATTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCTTACCTTCCATTTGATGACATACAGCTTCCCAAGATCTGTCTCCGACT 1380
QY 461 GlnArgArgGlyValArgThrGlyArgGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACGTCCGGGAGAGACTGGCAGGGGAGAGCCAGGATCATCAAGATTTGTGGACCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCCGCGCATGTTCAGCTGTCCGCTCTGTGAGGTATGACGAGGCTGT 1500
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGTATGAGCTTACCGCCCGCCAGACTACATTAAGCTACGACGTCATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGlnGlyValPheThrGlyLeu 540
DB 1561 CCGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACTCATATGATGATGCCCATTTCTATCCAGACAAACAGAGTGGGAGAACCTTCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
DB 1681 CTGTAGCTTACCAACCAACCGTGTGCTAGGGCTCAAGACCTTCCCATCTGCGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGAAGTGTGATTTGATGCTCCCAAGCCCACTCATAGGGCAACCCCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIleTyrIle 620
DB 1801 TACAGACTGGGCGCTGTGATGAATGAATACACCTGACACCCAGTCCAAATACATC 1860
QY 621 MetThrCysMetSerIleAspLeuGluValIleThrSerThrTyrValLeuValGlyGly 640
DB 1861 ATGACATGATGTGGCCCACTCGAGGTGTCACAGCACTCGGTGCTGTTGGCGG 1920
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValAlaIleValGlyArg 660
DB 1921 GTCTGTGCTGTGGCCCGCTATTCCTGTCAACAGCTGCTGTCATAGTGGCAG 1980
QY 661 IleValLeuSerGlyIleProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1981 GTCTGTGTCCGGGAAGCCGCAATCATACCTGACAGGGAAGTCTTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
DB 2041 GATGAGATGAAGAAGTGC 2058

RESULT 13
US-09-881-654-1
; Sequence 1, Application US/09881654
; Patent No. 6532601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESK, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOSSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / P17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811

PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/280,867
PRIOR FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1
LENGTH: 2058
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence:
OTHER INFORMATION: representative NS3/4a conformational antigen
NAME/KEY: CDS
LOCATION: (1) .. (2058)
US-09-881-654-1

Alignment Scores:

Pred. No.:	0	Length:	2058
Score:	3565.00	Matches:	671
Percent Similarity:	99.13%	Conservative:	9
Best Local Similarity:	97.81%	Mismatches:	6
Query Match:	98.54%	Indels:	0
DB:	4	Gaps:	0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

```
QY 1 MetAlaProIleThrAlaYrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCATCAGGGGAGGAGCCGAGACAGAGGGGGCTCTTGGGTGATATATACC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 61 AGCTTAATGCGCCGGGCAAAAACCAAGTGGAGGTGAGTCCAGATTGTCAACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysThrThrValIleHisGlyAla 60
DB 121 GCCCAAACTTCCCTGGGAACTGATCATATGGGTGTGCTGACTGTCTACACAGGGGCC 180
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetCysThrAsnValAsp 80
DB 181 GGAAAGAGACATCGCTGCTACCCAGAGGCTCTTCTCATCATGATGTAACCAATGTAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTGTGGGCTGGCCGCTCCGAAAGTACGATGATGACACCTGCACTTGC 300
QY 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGCTCCCTCGGACCTTTACCTGTGACAGAGGACCGCATGTCTCCGTGCGCGCGG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerCysThrLeuLysGlySerSer 140
DB 361 GGTATATGACAGGGGAGGCTGCTGCTGCGCGCCATTTCTTCTTGAAGAGCTCTCG 420
QY 141 GlyIleProLeuLeuCysProIleGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGCGCGCGGGGACCGCTGGGACATATTAAAGGCGCGGTGTC 480
QY 161 ThrArgGlyValAlaIleValAlaAspPheIleProValGlySerLeuGlnThrThrMet 180
DB 481 ACCCGTGAAGTGGCTAAGCGGTGACTTATCCCTGTGGAGAACTTAAGAACACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerCysThrVal 200
DB 541 AGGTCCCGGTGTTCAAGGATTAATCTCTTCCACCAAGTAGTCCCGACAGCTTCAAGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaIleVal 220
DB 601 GCTACACCTCCATGCTCCACAGGAGGAGGCAAAAGACCAAGGTCCGGCTGCATATGCA 660
QY 221 AlaGlnGlyThrLysValLeuValLeuAsnProSerValAlaAlaIleThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATTAAGTGTAGTACTCAACCCCTCTGTGTGCAACACTGGGCTTTGGT 720
```

```
QY 241 AlaThrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAAAGGCTCATGGATCATCTTCAATCAGACAGCGGGGTGAACAAT 780
QY 261 ThrThrGlySerProIleThrIleThrIleThrIleThrIleThrIleThrIleThrIle 280
DB 781 ACCACGTGGAGCGCCCATCAAGTACTCACCCTACGGGCAAGTCTCTTCCAGCGGGGTGC 840
QY 281 SerGlyGlyAlaIleThrAspIleIleIleGlyAspGlyCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGGGCTTATACATTAATATTGTGACGATGCCATCCACGATGCAATCC 960
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyValAlaArgLeuThrVal 320
DB 901 ATCTGGGCAATGGCACTGCTGCTTGAACAGAGAGACTGGGGGGGAGACTGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTGCGCACCGCCACCCCTCGGGCTCGTCACTGTGCCCATCCCAACATCGAGAGGT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGlnAlaIle 360
DB 1021 GCTCTGTCCACACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCCAGATATATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGlyLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAGTGCACACATCGCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaIleThrThrArgGlyLeuAspValSerVal 400
DB 1141 AACTGTGTGCATTTGGGCATCAATGCGCGCTTACCTACCGCGGTGTAACGTGTCGTC 1200
QY 401 IleProThrSerGlyAspValValAlaIleThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGCCCATCGGCGATGTGTCTGTGTCGCAACGATGCCCTCATGACCGGCTATAG 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValIleThrValAlaAspPheSer 440
DB 1261 GCGCACTTGCATCGGTGATAGACTCAATACGTGTGTCCACGACAGTCCGATTTCCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCCCTACCTTCAACATTAAGACATCACTCCCGCCCAATGCTGTCTCCGACT 1380
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleIleThrArgPheValAlaProGly 480
DB 1381 CAACGTGCGGGGACGACTGGCAGGGGAGGAGCCAGCATCTACGATTTGTGGACCGGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysThrAspAlaGlyCys 500
DB 1441 GAGGCGCCCTCGCGCATGTTCGACTGCTGCTGTGTGATGATGACGACAGGCTGT 1500
QY 501 AlaTrpTrpGluLeuThrProAlaGluThrThrValArgLeuArgAlaIleThrMetLeuThr 520
DB 1501 GCTTGTATATAGCTACGCGCGCGGAGACTACATTAAGGTATGAGAGGTACATGAACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrIleGluValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCGTGTGCCAGACATCTTGAAITTTGGAGGGCGCTTTTACAGGCTTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyLeuAsnLeuProThr 560
DB 1621 ACTCATATATAGTCCCATCTTATCCCAAGACAAACAGAGTGGGAGAACTTCTTCCCTAC 1680
QY 561 LeuValAlaIleArgAlaIleThrValCysAlaIleArgAlaIleProProProSerTrpAsp 580
DB 1681 CTGTATGCGTATCAAGGACCGGTGTGCTAAGGAGCTCAACCCCTCCCATGTGGAGAC 1740
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGAAGTGTGATTTGATTCGCTCAGCCACCTCATATGGGCCAACACCCCTGTCTA 1800
```

QY 601 TyrArgLeuGluValAlaGlnAsnGluValThrLeuThrHisProValThrIleValGlyTyrIle 620
DB 1801 TACAGACTGGCGCCCTGTTCCAGAAATGAATCACTCCAGACCCAGTCCACCAATATCATC 1860
QY 621 MetThrCysMetSerLysAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
DB 1861 ATACAGCATGTCGCGCCGACCTGAGAGTGTACAGACACCTGGGTGCTGTGGGGGC 1920
QY 641 ValLeuAlaAlaLeuAlaAlaValTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 1921 GTCTGCTGCTCTTGGCGCGGTATGCTCTCAAGGCTGCGTGTCAATAGTGGGAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1981 GTGCTCTTGTCCGGGAAGCCGCAATCATCTGACAGGAGAGTCTCTACCAAGATTTC 2040
QY 681 AspGluMetGluGluCys 686
DB 2041 GATGAGATGAGAGAGTGC 2058
RESULT 14
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 615087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBY Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8985
; US-08-444-818-137
Alignment Scores:
Pred. No.: 0 Length: 8987
Score: 3565.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.54% Indels: 0

DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)
QY 1 MetAlaProIleThrAlaTyrAlaGlnIleThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3076 CTGGCCCATCATCAGCGGATGCGCCAGACAGAGAGGGGCTCTTAGGGTGCATATCACC 3135
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 3136 AGCTTAACCTGGCCGGGACAAAACCAAGTGAAGGTGAAGTCCAGATTGTGTAACCTGCT 3195
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValIleGlyAla 60
DB 3196 GCCCAAACTCTCTGCAACAGTGCATCAATGAGGTGTGCTGCACTGTCAACACGGGGCC 3255
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGluMetTyrThrAsnValAsp 80
DB 3256 GGAAACGAGACCATCCGTCACCCCAAGGATCTGTATCCAGATGTAACCAATGTAGAC 3315
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3316 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGTGAAGCCGCTCATATGACACCTGCATTCG 3375
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 3376 GGCTCTCGGACCTTACCTGATCAGAGGACAGCCGATGTCACTCCGTGCGCGCGCG 3435
QY 121 GlyAspGlyArgGlySerLeuLeuSerProAlaArgProIleSerTyrLeuLysGlySerSer 140
DB 3436 GGTGATAGCAGGGGCGACCTGTGTCGCCCGCCCATTTCTCATCTTGAAGGCTCTCG 3495
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgHisAlaValCys 160
DB 3496 GGGGATCCGCTGTGTGTCGCCCGGGGACGCGGTGGCATTTAGGCGCGCGGTGTGC 3555
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGlnThrMet 180
DB 3556 ACCCGTGAAGTGGCTAAGCGCGGTGATCTTATCCCTGTGAAGACCTTAGACCAACCATG 3615
QY 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
DB 3616 AGGTCCCGCGTGTTCACGATTAACCTCTCCACAGTGTGTCGACGAGCTTCGAGG 3675
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrIleValProAlaIleTyrAla 220
DB 3676 GCTCACCTCATGCTCCACAGGACGGGCAAAAGCACCAAGGTCCGGCTGCATATGCA 3735
QY 221 AlaGlnGlyTyrLeuValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 3736 GCTCAGGGCTATTAAGTGTGTAATCAACCTCTGTGTGTCGACACACTGGGCTTTGGT 3795
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 3796 GCTTACATGTCCAAAGGCTCATGGGATCGATCTTAACATCAGAGCCGGGTGAGAACAAAT 3855
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 3856 ACCACTGGCAGCCCATCAAGTCACTCCACTCAAGGAGGTCTCTTGGCCACGGCGGGTGC 3915
QY 281 SerGlyValAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 3916 TCGGGGGGCGCTTATATCAATATTAATTGTGACAGTGCACCTCCACGATGCCACATCC 3975
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 3976 ATCTGGGATGGGACGCTGCTTGCACAGCAGAGACTCGGGGGGCGAGACTGTTGGG 4035
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4036 CTGCGCACCGCACCCCTCCGGGCTCCGTCATGTGCCCATTCACCAATCGAGGGGTT 4095
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyAlaIleProLeuGlnAlaIle 360

```

Db      4096 GCTCTGCCACCAACCGAGAGATCCCTTTTACGGCAGGCTATCCCTCCGAATATC 4155
Qy      361 LysGlyValArgHisLeuIlePheCysHisSerIleValCysAspGluLeuAla 380
Db      4156 AAGGGGGGAGACATCTCATCTTCTGTCATCAAGAAAGAGCGCAGCATCGCGCA 4215
Qy      381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      4216 AAGCTGTCCCATTTGGGCATCAATGCCGTGCTACTACCGCGGTCTTGAAGTCCGTC 4275
Qy      401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      4276 ATCCCGACCAAGCGCGATGTTGTCGTGCGCAACCATGCCCCCTCATGACCGGCTATACC 4335
Qy      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      4336 GGGCACTTCCACTCGGATGATGACTGCATACGTTGTCAACCCAGACAGTGCATTCAGC 4395
Qy      441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db      4396 CTGACCTTACCTTCAACATTTAGACATACGCTCCCGCAGAGATGCTGCTCCGCGACT 4455
Qy      461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyTyrTyrArgPheValAlaProGly 480
Db      4456 CAACGTGGGGAGAGACTGGCAGGGGAGAGCAGGACATCAACAGATTGTGGCAGCGGG 4515
Qy      481 GlnArgProSerGlyMetPheAspSerSerValLeuGlyCysTyrAspAlaGlyCys 500
Db      4516 GAGGCCCCCTCCGACATGTTGACTGCTCGCTCTGTAGAGCTATGACGAGGCTGT 4575
Qy      501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db      4576 GCTTGGATATAGCTCAAGCCCGCAGACATCAAGTTAGGCTACGAGGCTACAGAAACCC 4635
Qy      521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db      4636 CCGGGGCTTCCCGGTGTCAGACCATCTGAATTTGGAGGGCGCTTTTACAGGCGCTC 4695
Qy      541 ThrHisIleAspAlaHisPheLeuSerGlnThrIleGlnSerGlyValAsnLeuProTyr 560
Db      4696 ACTCATATAGATGCCCTTCTTATCCAGCAAGAGAGAGTGGGGAACCTTCTTAC 4755
Qy      561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db      4756 CTGGTACGTCACCAAGCCACCGTGTGCGTACAGGCTCAAGCCCTCCCATGCTGGGAC 4815
Qy      581 GlnMetTrpLysCysLeuIleArgLeuLysProThrIleuHisGlyProThrProLeuLeu 600
Db      4816 CAGATGTGGAAGTGTGATTCGCTCAAGCCCACTTCATGGGCCCAACACCCCTGCTA 4875
Qy      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      4876 TACGACTCGGCGCTGTTCCAGATGAATACCTCTGACGACCCAGACCAACCAATATAC 4935
Qy      621 MetThrCysMetSerAlaAspLeuGluValValIleThrSerThrTrpValLeuValGly 640
Db      4936 ATGACATGCAATGTGCGCGACCTGAGAGTGTGTACGACACTGGGAGTCTGTTGGGCG 4995
Qy      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      4996 GTCTGCTGCTGTTGGCGCGCTATTTGCTGTCAACAGGCTGCGGTCAATAGTGGGAGG 5055
Qy      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      5056 GTCTCTTGTCCGGGAAGCGGCAATACATCTGACAGGAAGTCTTACCGAGAGTTC 5115
Qy      681 AspGluMetGluGluCys 686
Db      5116 GATGATGAGAGAGTGC 5133

```

RESULT 15
US-08-444-818-53

```

; Sequence 53, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; TITLE OF INVENTION: NANOV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..5360
; US-08-444-818-53

Alignment Scores:
Pred. No.: 0 Length: 5360
Score: 3555.00 Matches: 668
Percent Similarity: 99.428 Conservative: 13
Best Local Similarity: 97.524 Mismatches: 4
Query Match: 98.264 Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-53 (1-5360)
Qy      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      930 CTGGCGCCCATCAAGCGGTCATACGCCAGACAGACAGAGGGCGCTCTTGGGTGATATACCC 989
Qy      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db      990 AGCTTAACCTGGCGGAGCAAAAACCAAGTGAAGGTGAAGTCCAGATTGTGTCACTGCT 1049
Qy      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db      1050 GCCCAACCTTCTCTGCAACGTGCATCATGCGGTGCTGCACTGTACACACGGGGCC 1109
Qy      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      1110 GGAACGAGGACATTCGCTCACCCAGAGGGTCTGTGTCATCAGATGATATCAATGTAGAC 1169
Qy      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

```

Db 1170 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGTAGACCGCTCATGTGACACCCCTGCACCTTGC 1229
 Qy 101 G1ySerSeraspLeuTyrLeuVal1ThrArgH1sa1aaspVal1IleProVal1ArgArgArg 120
 Db 1230 GGGTCTCTGGACCTTAACTGTGTACAGAGCAGCCGATGTCTATTCCTGCGCGCGG 1289
 Qy 121 G1yAspG1yArgG1ySerLeuLeuSerProArgPro1IleSerTyrLeuYsg1ySerSer 140
 Db 1290 GGTGATAGCAGGGGAGCTGTGTGCCCCGGCCATTTCTTAATTGAAAGGCTCTCG 1349
 Qy 141 G1yG1yProLeuLeuCyProAlaG1yH1a1aValG1y1IlePheArgAla1aValCys 160
 Db 1350 GGGGGTCCGCTGTGTGTGCCCCGGGGCAGCCGTGGCATATTATTTAGGGCCGCGGTGTGC 1409
 Qy 161 ThrArgG1yVal1AlaYsa1aValaaspPhe1IleProVal1G1ySerLeuG1uThrThrMet 180
 Db 1410 ACCGGTAGAGTGGCTTAAGCGGTGACTTATTCCTGTGAGAACTTAAGACAAACCATG 1469
 Qy 181 ArgSerProVal1PheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1470 AGGTCCCGGTGTACAGGATACTCTCTCCACAGATAGTGCCCGCAGACTTCCAGGTG 1529
 Qy 201 AlaH1sLeuH1sa1aProThrG1ySerG1yYsSerThrYsVal1ProAla1a1yYrAla 220
 Db 1530 GCTCACCTCCATGCTCCACAGGACGCGCAAAAGCACAAAGTCCCGCTGCATATGCA 1589
 Qy 221 AlaG1yG1yTyrYsVal1LeuVal1LeuVal1ProSerVal1Ala1aThrMetG1yPheG1y 240
 Db 1590 GCTCAGGGCTATTAAGGCTAGTACTAACCCCTGTGTCTGTGACAACTGGGCTTTGGT 1649
 Qy 241 AlaTyrMetSerYsAlaH1sG1y1IleAspProHsn1IleArgThrG1yVal1ArgThr1Ile 260
 Db 1650 GCTTACATGTCCAAGGCTCATGGATGATCTTAACATCAGAACCGGGGAGAGAACATTT 1709
 Qy 261 ThrThrG1ySerPro1IleThrTyrSerThrTyrG1yYsPheLeuAlaaspG1yG1yCys 280
 Db 1710 ACCACTGGCAGCCCATCAGTACTCCACCTACGCGCAAGTTCCTTGGCGGCGGTGC 1769
 Qy 281 SerG1yG1yAlaTyrAsp1Ile1IleCysAspG1uCyH1sSerThrAspAla1aThrSer 300
 Db 1770 TCGGGGGGCGCTTATGACATTAATTGTGACAGAGCCACTCCACGAGATGCCACATCC 1829
 Qy 301 1IleLeuG1y1IleG1yThrVal1LeuAspG1n1aG1uThrAlaG1yAla1aArgLeuThrVal 320
 Db 1830 ATCTTGGGCAATCGGCATGTCTCTTGAACAGAGACTCGGGGGGAGATCGTGTGTG 1889
 Qy 321 LeuAla1aThrAla1aThrProProG1ySerVal1ThrVal1ProH1sProAsn1IleG1uG1uVal 340
 Db 1890 CTCGCCACCGCCACCCCTCCGGGCTCGTCACTGTGCCCATCCCAATCGAGAGGTT 1949
 Qy 341 AlaLeuSerThrThrG1yG1u1IleProPheTyrG1yYsa1a1IleProLeuG1uAla1Ile 360
 Db 1950 GCTCTGTGCCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTATATC 2009
 Qy 361 YsG1yG1yArgH1sLeu1IlePheCyH1sSerYsYsYsYsCyAspG1uLeuAla1a 380
 Db 2010 AAGGGGGGAGACATCTCATCTTCTGTCTATTCAAAGAAAGATCGACGAACTCGCCGCA 2069
 Qy 381 YsLeuVal1a1aLeuG1yVal1aAsn1aVala1aTyrTyrArgG1yLeuAspVal1SerVal 400
 Db 2070 AAGCTGGTCCATTTGGGCAATGCGGTGCTACTACCGGGTCTTGAAGCTGTCCGTG 2129
 Qy 401 1IleProThrSerG1yAspVal1a1aVal1a1aThrAspAlaLeuMetThrG1yPheThr 420
 Db 2130 ATCCCGACGAGCGCGAGTGTGTGTGTGTGCAACCGATGCCCTCATGACCGGCTATATCC 2189
 Qy 421 G1yAspPheAspSerVal1IleAspCyAsnThrCyVal1ThrGlnThrVal1aaspPheSer 440
 Db 2190 GGGCACTTCACTCGGATGACTGCAATACGTGTCTACCCAGACAGTCAATTTTCAAGC 2249
 Qy 441 LeuAspProThrPheThr1IleG1uThr1IleThrLeuProGlnAspAlaVal1SerArgThr 460
 Db 2250 CTTGACCTTACCTTCACTTGAAGACATACAGCTCCCGAGATGCTGTCTCCGCACT 2309

Qy 461 GlnArgArgG1yArgThrG1yArgG1yYsProG1y1IleTyrArgPheVala1aProG1y 480
 Db 2310 CAACGTGGGGGAGAGACTGGCAGGGGGAAAGCCAGGATCTACAGATTTGTGGACCGGGG 2369
 Qy 481 GlnArgProSerG1yMetPheAspSerSerVal1LeuCyG1uCyTyrArgAla1aG1yCys 500
 Db 2370 GAGCGCCCTCCCGGCATGTGTGCACTGTCCCTCTGTGAGGCTATGACGAGGCTGT 2429
 Qy 501 AlaTrpDyG1uLeuThrProAlaG1uThrThrVal1ArgLeuArgAla1aTyrMetAsnThr 520
 Db 2430 GCTTGTATGAGCTCAAGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 2489
 Qy 521 ProG1yLeuProVal1CyG1nAspH1sLeuG1uPheTrpG1uG1yVal1PheThrG1yLeu 540
 Db 2490 CCGGGGCTTCCCGTGTGCCAGACATCTTGAATTTTGGAGGGCGCTTTTACAGGCTTC 2549
 Qy 541 ThrH1s1IleAspAlaH1sPheLeuSerGlnThrYsGlnSerG1yG1uAsnLeuProTyr 560
 Db 2550 ACTCATATAGATGCCCATTTCTATCCAGACAAACAGAGTGGGAGAACTTCTCTTAC 2609
 Qy 561 LeuVal1a1aTyrG1nAla1aThrVal1CyAlaArgAlaG1nAla1aProProProSerTrpAsp 580
 Db 2610 CTGGTAGCTTACCAAGCCACCGGTGCTAGGGCTCAAGCCCTCCCATGTGGAGAC 2869
 Qy 581 GlnMetTrpYsCySerLeu1IleArgLeuYsProThrLeuH1sG1yProThrProLeuLeu 600
 Db 2670 CAGATGTGAAGTGTGATTTGCTCTCAAGCCCAACCTCATGGGCAACCCCTGTCTA 2729
 Qy 601 TyrArgLeuG1yAlaValaGlnAsnG1uVal1ThrLeuThrH1sProVal1ThrYsTyr1Ile 620
 Db 2730 TACAGACTGGGCGCTGTTCAGAAATGAATCAACCTGACGACCACTCAACAAATATCATC 2789
 Qy 621 MetThrCyMetSerAlaAspLeuG1uVal1a1aThrSerThrTyrVal1LeuVal1G1yG1y 640
 Db 2790 ATGACATGATGTGCGGCCGACCTGAGGTGTCTACAGACACTGGGTGTCTGTGTGCGGC 2849
 Qy 641 ValLeuAla1aLeuAla1aTyrCyLeuSerThrG1yCyAsVal1a1IleVal1G1yArg 660
 Db 2850 GTCTGTCTGTCTTGGCGCGGATTTGCTGTCAACAGGCTGCTCATAGTGGGCAAG 2909
 Qy 661 1IleValLeuSerG1yYsProAla1a1IleProAspArgG1uVal1LeuTyrArgG1uPhe 680
 Db 2910 GTGCTCTTGTGCGGAAAGCGGCAATCATACCTGACAGGGAAGTCTCTTACGAGAGTTC 2969
 Qy 681 AspG1uMetG1uG1u 685
 Db 2970 GATGAGATGGAAGAG 2984

Search completed: September 17, 2004, 10:32:31
 Job time : 214 secs

THIS PAGE BLANK (USPTO)